

# Can Bonus Packs Mislead Consumers? A Demonstration of How Behavioural Consumer Research Can Inform Unfair Commercial Practices Law on the Example of the ECJ's Mars Judgement

Kai P. Purnhagen<sup>1</sup> · Erica van Herpen<sup>2</sup>

Received: 21 March 2016 / Accepted: 27 March 2017  
© The Author(s) 2017. This article is an open access publication

**Abstract** The use of psychological findings in EU internal market regulation has gained interest, particularly in the area of unfair commercial practices. This study investigates consumer perceptions of bonus packs containing an oversized indication of the “extra” volume in the package, such as in the Mars case. The Mars case serves as a standard reference in EU unfair commercial practices law which is used as a benchmark to determine the “average consumer.” Our study demonstrates how an experiment can be set up to provide empirically based insights on whether a practice is “deceptive.” Results of our experiment show that consumers overestimate the extra volume when confronted with an oversized indication compared to control conditions, which is first empirical evidence that this practice is potentially deceptive.

**Keywords** Behavioural consumer · Unfair commercial practices · Average consumer · Mars judgement

---

✉ Kai P. Purnhagen  
kai.purnhagen@wur.nl

Erica van Herpen  
Erica.vanherpen@wur.nl

<sup>1</sup> Law and Governance Group, Erasmus University of Rotterdam, Law Faculty, Wageningen University, Hollandseweg 1, 6706 KN Wageningen, The Netherlands

<sup>2</sup> Marketing and Consumer Behaviour Group, Wageningen University, Hollandseweg 1, 6706 KN Wageningen, The Netherlands

The potential use of psychological findings in EU internal market regulation is gaining increased interest in recent years. Scholars (Amir and Lobel 2012; Burgess 2012; Franck and Purnhagen 2014; Purnhagen 2015; Purnhagen and Feindt 2015; Sibony 2015; Sibony and Alemanno 2015; Trzaskowki 2011; van Aaken 2015) as well as members of the highly influential Joint Research Centre of the European Commission (van Bavel et al. 2013) have asked for more recognition of behavioural studies in EU internal market regulation, particularly in the area of unfair commercial practices (Duivenvoorde 2015; Purnhagen 2017; Sibony 2015; Trzaskowki 2011; critical towards inclusion of such studies in unfair commercial practices law: Weatherill 2007, pp. 128–129). In this respect, in a recent study for the European Commission, behavioural research was extensively applied to Unfair Commercial Practices Law.<sup>1</sup> Although the fact that behavioural studies can be relevant has thus been argued (see for further academic treatments of these issues the contributors to Zamir and Teichman 2014 and Mathis 2015), there is little guidance so far on *how* such studies can be conducted to inform policy makers and regulators.

The current paper illustrates the potential of behavioural science insights to inform EU unfair commercial practices law by first applying a behavioural science model to the conflict situation in the Mars case (2). In this case, the addition of extra volume for the product was indicated on pack, in a coloured area that was substantially larger than the extra volume. The ensuing question was whether consumers are deceived by the size of the coloured area, even though the correct percentage of extra volume was indicated as well. Subsequently, we will test the effects of an oversized “extra volume” area empirically in a simulated setting similar to the Mars case. Specifically, we will provide a procedure to test consumer perceptions of a bonus package containing an oversized indication on the extra volume in the package (e.g., an indication of 10% extra on a coloured area on the package that is considerably larger than 10%, as in the Mars case) to determine the extent to which such a label is deceptive to the consumer (3) and an initial empirical demonstration (4). The current study thus focuses on deceptiveness of the label in terms of product perceptions. Whether this in turn influences purchase decisions (which would be considered misleading<sup>2</sup>) and the concrete potential effects on law, regulation, and policy making are topics that are left for future research. The empirical study is a small-scale experiment using student participants, revealing the types of insights that could be gained by such an approach. The experiment can easily be upscaled to other countries and more representative samples in future research, to increase external validity. Yet, given that the underlying phenomenon is a very robust and often demonstrated bias (anchoring and adjustment), it is likely that the results of our experiment are replicable and can be relied upon (5). We conclude that the behavioural model presented has the potential to serve as a robust model for decision making in EU law, which provides relevant insights that can be taken into deliberation (6). That is to say, that we do not advocate that regulation needs to be based on empirics or even that empirical findings would trump other factors, but rather that behavioural research can provide insights which could be taken into consideration at Member State or EU level. Empirical data could be taken as an indication for a legislative, regulatory or judicial decision, weighed on its merits (e.g., how convincing are the data, how well was the study conducted (see further on the need for good data for good policy making Purnhagen and Feindt 2015), integrated into the regulatory process, and, e.g., used as an argument within the exercise

<sup>1</sup> Consumer vulnerability across key markets in the European Union, Final Report, January 2016.

<sup>2</sup> See for the requirements for both tests to be positive to determine “misleading” Art. 6 (1) UCPD.

of the proportionality test of infringements of the free movement of goods (see further in this respect Purnhagen 2017).

## The Mars Case and the Anchoring Effect

In this section, we will first introduce the facts of the Mars case (1). Subsequently, we illustrate how this case could well be understood as dealing with the “anchoring” bias known from behavioural research (2).

### 1. The facts of the Mars case

The case underlying our research concerns a standard case of EU unfair competition law.<sup>3</sup> It has been described as “one of the clearest and most specific examples of the expectations of the CJEU (Court of Justice of the European Union, addendum by authors) towards consumers.” (Duivenvoorde 2015, p. 40) Indeed, the Mars case serves as a general reference to illustrate the benchmark consumer in the Court’s case law on the free-movement rules and consumer protection (Stuyck 2010, p. 14). This case is hence a good illustration of how the normative concept of the average consumer compares to the “real world” of consumers as they behave in an experimental setting. It also illustrates how a potential change of the consumer concept in marketing law towards a benchmark that incorporates insights from behavioural science may affect the outcome of judgments in the area.

The facts of the case are simple: In 1993, the Mars company had launched a European-wide marketing campaign selling a packaged ice cream bar with 10% more content for the regular price. Mars indicated the increase of 10% on the product packaging. However, the underlying coloured area of the 10% increase announcement took up nearly 30% of the total surface area of the wrapping. Initially, it was on a German court to rule whether this practice was “deceptive” according to German law. As this question touched upon the interpretation of EU law also, it stayed proceedings and referred questions to the European Court of Justice (ECJ). The Court had to answer the question whether this oversized area would induce consumers into believing that the increase was larger than the actual 10% increase and therefore would deceive them. The Court answered in the negative by relying on the normative notion of a “reasonably circumspect consumer” as the benchmark.<sup>4</sup>

In introducing behavioural insights into the notion of the internal market in general and the area of unfair commercial practices in particular, the Mars case may be the ideal case to start with, for several reasons. First, the case is well known and has served as a benchmark to illustrate the concept of the “average consumer” and thus has been influential in guiding other cases (Stuyck 2010, p. 14)<sup>5</sup> and legal acts. Using this case as an illustration for the potential use of psychological insights based on empirical tests could thus be likewise influential. Second, the case is specific and allows for a straightforward application in an experimental setting. Third, the practice under

---

<sup>3</sup> *Verein gegen Unwesen in Handel und Gewerbe Köln e.V. v Mars GmbH* (C-470/93) [1995] E.C.R. I-1923.

<sup>4</sup> *Verein gegen Unwesen in Handel und Gewerbe Köln e.V. v Mars GmbH* (C-470/93) [1995] E.C.R. I-1923, para 24.

<sup>5</sup> The so-developed concept of the “average consumer” was even transferred to other areas of internal market law such as the area of unfair contract terms (see *Árpád Kásler, Hajnalka Káslerné Rábai v OTP Jelzálogbank Zrt* (C-26/13), ECLI:EU:C:2014:282).

investigation resembles situations in which a systematic bias, the anchoring effect, has been shown to occur. This is one of the first biases uncovered in psychology (Thaler and Sunstein 2008) and remarkably robust (Furnham and Boo 2011; Mussweiler et al. 2000). Thus, although the case at hand has not been put under empirical test so far, the similarities with the anchoring effect allow for the formulation of specific expectations that can be tested in an empirical study. Following Schuck (1989, p. 323), empirical studies refer here to “statistical studies, i.e., those that involve the application of statistical techniques of inference to [..., omission by authors] data in an effort to detect important regularities (or irregularities).” Although this definition may be considered narrow, as it excludes case studies, legal history, and other legal studies that can also be considered empirical, it has the advantage of focusing clearly on one type of legal scholarship and distinguishing it from other approaches (Schuck 1989; Heise 1999).

## 2. The anchoring effect in the Mars case

When people are confronted with a bonus pack, their estimates of the amount of extra volume that is offered can depend on other (relevant or irrelevant) values that are activated. The coloured area on the pack offers such a value. The process by which consumers are biased in their estimates of values because a sense of size has been activated and influences these estimates has been examined in anchoring literature.

Although early mentioning of anchoring had already occurred in psychophysics, the effect gained more widespread recognition by the influential work of Tversky and Kahneman (1974). The notion of the anchoring effect is that people often make estimates by starting from an initial value (the “anchor”), with estimates being biased towards this initial value. Different underlying processes have been proposed for the anchoring effect (see Furnham and Boo 2011 for a review), including insufficient adjustments from the anchor, attempts of people to confirm that the anchor value is the correct answer, and the use of anchors as a cue to a reasonable answer.

The anchoring effect is one of the most robust biases that have been shown to exist in decision making (Furnham and Boo 2011). It occurs even when anchor values are clearly uninformative, appears to be independent of a person’s motivation to examine the issue at hand, and is not mitigated by explicit instructions to correct for the potential influence of the anchor (Mussweiler et al. 2000). It has been shown across various domains, including the legal domain. For instance, when a cap placed on punitive damages is high, it can serve as an anchor such that the size of awarded punitive damages increases as well (Robbennolt and Studebaker 1999). Likewise, sentencing demands of prosecutors influence sentencing decisions (Englich 2006).

Anchors can come in many forms, which need not be numerical, and anchoring effects may also occur across modalities (Oppenheimer et al. 2008). That is, anchors may activate a sense of size, which is not attached to any particular rating scale, and may bias subsequent judgments (LeBoeuf and Shafir 2006; Oppenheimer et al. 2008). Furthermore, research on the effect of multiple anchors has shown that people remain susceptible to (unreliable) anchors even in the context of other (more relevant) anchor values (Whyte and Sebenius 1997).

In the Mars case, the bonus pack contains two potential anchors for the amount of extra volume that the package contains: the percentage provided on pack (a correct anchor) and the coloured part of the package area (an incorrect anchor). Given that prior research has indicated that judgments are influenced by unreliable anchors even in the presence of more reliable anchors (Whyte and Sebenius 1997) and the robustness of the anchoring phenomenon, we expect that consumers will be biased by the incorrect anchor in this case as well. This is further based on the difference in modality between the two anchors. Whereas the percentage is numerical and needs to be translated into a volume estimate to be meaningful for consumers, the coloured area directly indicates a part of the packaging. Although anchoring effects can occur across modalities (Oppenheimer et al. 2008), anchors in the same modality should be easier to use for people. Thus, we expect that people's estimates of the extra volume that a bonus pack contains are influenced by the coloured area on the pack, such that an oversized area will deceive consumers into thinking that the package contains more extra volume than is actually present.

Prior studies in the marketing and consumer behaviour literature have examined consumer response to bonus packs in general, often in comparison to price discounts (Hardesty and Bearden 2003; Xu and Huang 2014). This stream of literature has indicated a preference of consumers for bonus packs over price discounts for virtue products (Mishra and Mishra 2011) and for smaller-sized promotions (Diamond 1992). More relevant to the current investigation, consumers have been shown to have difficulties in processing numeric information such as percentages (Chen and Rao 2007), which leads them to neglect base values, and prefer bonus packs over economically equivalent price promotions (Chen et al. 2012). Given this difficulty in interpreting percentages, consumers may be more susceptible to influences of a coloured area on pack. In another stream of research, consumer perceptions have been shown to depend on the spatial dimensionality of product packages (Chandon and Ordabayeva 2009; Wansink and van Ittersum 2003). That is, consumers perceive tall and narrow containers differently from short and wide containers, even when objectively their volume is the same. Yet, to our knowledge, no empirical study has so far examined how a coloured area on a product's packaging affects proportional volume perceptions for bonus packs. Therefore, we set up a study to demonstrate how this specific issue can be empirically tested and to provide a first test.

To determine the effects of the oversized coloured area as well as of the actual percentage that was provided on the pack, the study used several comparison conditions in an experimental design, as will be explained in the "Method" section. Because the specific Mars product that inspired the original Court ruling is no longer available on the market, we used a similar product that we found in a Dutch supermarket. This is a bonus package of coffee, which shares the specific elements that are also present in the Mars case: a verbal (correct) indication of 10% extra presented in a larger coloured area on pack (28% of the wrapper). If our expectation that the coloured area on pack can deceive consumers is supported by the data, this provides important insights which can be of relevance to legislators, courts and regulators when exercising their discretion or interpreting the law.

### 3. A procedure to test "deceptiveness" in the Mars case

The anchoring effect typically occurs without people being aware of it. It would

therefore be futile to ask people to self-report on the effects of the bonus area. They may not be aware of being deceived, and social desirability may also bias their answers. Therefore, we use an experiment to causally determine if the coloured bonus area leads to an overestimation of the extra volume that is provided.

In addition to comparing an experimental condition with an oversized bonus area to a control condition with a correctly sized bonus area, we include two additional conditions to gain more insights. The first is a condition with only the text percentage and no coloured area on pack (a “percentage-only” condition), and the second is a condition with only the incorrect coloured area and no text percentage (an “area-only” condition). These conditions can help determine the extent to which people (incorrectly) use the coloured area on the pack in their volume estimates and the extent to which they use the (correct) percentage provided. Using a design with multiple comparison conditions allows us to disentangle several possibilities. Specifically, if participants are able to correctly perceive the amount of bonus volume based on the percentage indicated on the pack and are not influenced by the oversized area at all, perceptions of the “oversized area” pack and the “correct area” pack should be similar. In contrast, if participants base their perception of the amount of bonus volume on the indicated area and disregard the percentage that is indicated, perceptions of the “oversized pack” should be similar to those of the area-only pack (where no percentage is indicated). Finally, if participants fail in using the percentage information correctly, without being affected by the oversized area per se, perceptions of the oversized pack should be similar to those of the percentage-only pack.

The use of this set of comparison conditions also allows us to investigate the possibility that there is an issue of statistical power in the experiment. That is, there should be differences in people’s volume perceptions between the area-only and the percentage-only conditions, as these provide only one anchor of either 10 or 28%. Should these two conditions not significantly differ from each other, this could point towards a power issue (i.e., a small magnitude of the effect and/or a too small sample size) and other results should be interpreted with great care.

#### 4. Putting the Mars case to the empirical test

The objectives of the empirical study are to demonstrate, on a small scale, what type of insights can be gained from an experiment such as described in the “[Discussion](#)” section. Because the main purpose is demonstration of the method, the sample will be a non-representative convenience sample. Replication of the experiment using a representative sample or in other countries or situations is relatively straightforward.

## Method

**Participants and Design of the Experiment** Participants were 126 students of a Dutch university (35% male and 65% female; mean age 22 years). They were randomly assigned to one of four conditions. These four conditions differed in the type of package that participants saw, as introduced in the “[Discussion](#)” section. One group of participants saw a coffee package with caffeine-free coffee that contained 10% extra volume (550 g instead of 500 g), with a coloured yellow area on the package. The coloured area amounted to 28% of the (original) package volume, which is thus an oversized area compared to the actual additional volume in

the bonus pack. The package also contained a marking “+10% free” within this area. This package was available on the Dutch market at the time that the study took place (January 2014).

To examine whether the indicated area influences consumer perceptions, three other conditions were created by manipulating the packaging using Photoshop. Pictures of the packages are provided in [Appendix](#). One condition contained the same marking of +10% free but now in a yellow area that actually amounted to 10% of the package volume. This is thus a situation in which both anchors are consistent and correct. A second condition contained only the +10% free marking, and a final condition contained the wording extra volume in a yellow area of 28% of the original package volume. In these cases, only one of the two anchors was thus present. Both the original package and the adjusted packages were printed using a colour printer and wrapped around the bonus package. This was done to ensure that all packages (including the original one) had the same texture and quality of packaging. [Table 1](#) provides a description of the packages that were used in this experiment.

**Procedure** Participants were recruited by e-mail and flyers around campus. Upon entering the research room and giving informed consent, they were seated at a desk and presented with one of the four packages used in the experiment. Desks were separated by screens, so that participants could not see each other or the packages that other participants received. They were asked to answer eight general statements about the package, in order for them to pay attention to this package (e.g., “It is easy to notice from the pack that the coffee is caffeine free,” “The opening instruction on the side of the package is easy to understand,” “The text on the front of the package is easy to read,” “I like the design of the package;” all on five-point scales ranging from completely disagree to completely agree). None of these questions specifically focused on the bonus volume.

Next, the package of coffee was taken away, and participants answered a small filler task in which they indicated their agreement to five statements about the university campus (e.g., “I would appreciate the addition of a small supermarket on campus”). This filler task was included to ensure that participants could no longer see the package when answering the next set of questions about product perceptions (see “[Measures of Volume Perception](#)” section). After answering these questions on product perceptions, participants indicated their gender, age, and study domain. They were thanked and received a snack product of their choice as a token of appreciation (candy bar, apple, or crisps).

**Measures of Volume Perception** Prior research has shown that the measurement instrument used to measure perceptions should be carefully constructed as instruments differ in their sensitivity in picking up differences (Price et al. 1994), and measures of memory (for example, recall of a specific number) may not always reveal what consumers actually know (Vanhuele and Drèze 2002). Therefore, we used three different ways to measure volume perceptions. In addition to asking a numerical estimate of the percentage of bonus volume

**Table 1** Description of the four different packages used in the experiment

Condition	Size of yellow area	Text on package
Oversized area	28%	10% free
Correct area	10%	10% free
Percentage only	No yellow area indicated	10% free
Area only	28%	Extra volume

and using traditional self-report scales, we asked people to visually indicate the bonus volume (akin to a two-dimensional version of the visual analogue scale). The three measures are as follows:

1. Visual estimate: participants were shown the contours of the bonus package and asked to indicate the amount of bonus volume that was present in the pack.
2. Perception of bonus volume and pack size (self-report scales): participants answered questions on a seven-point scale ranging from completely disagree to completely agree. Perception of bonus volume was measured with one item (“This coffee pack offered me a lot of bonus volume”), and perception of pack size was measured with two items (“This coffee pack was a large pack” and “This coffee pack was much larger than most coffee packs”). As a reliability analysis provided a satisfactory Cronbach’s alpha of 0.69, the mean of these two latter items was used in subsequent analyses.
3. Estimated volume percentage: participants indicated the (numerical) percentage of the bonus volume. They were asked to give an estimate when they did not know the answer.

**Analyses** Univariate analyses of variance (ANOVA) were used to analyse the data. Separate analyses were conducted for each of the measures of volume perception, and condition was the independent variable in each of these analyses. The output of these analyses (in the form of an  $F$  test) indicated whether the means of the four conditions were significantly different from each other (at  $p < .05$ ), for each of the measures of volume perception. When the ANOVA indicated a significant difference, we used the least significant difference (LSD) pairwise multiple comparison test to investigate which pairs of conditions were significantly different. This is equivalent to performing multiple individual  $t$  tests between all pairs of groups.

In addition to these tests, we also specifically tested the area that participants indicated was the bonus volume against the objectively correct area. This latter area was set at a boundary of 20 mm from the top of the package (upwardly rounded number). One-sample  $t$  tests were used for each condition separately, to test if reported areas significantly differed from the actual bonus volume in the pack.

## Results

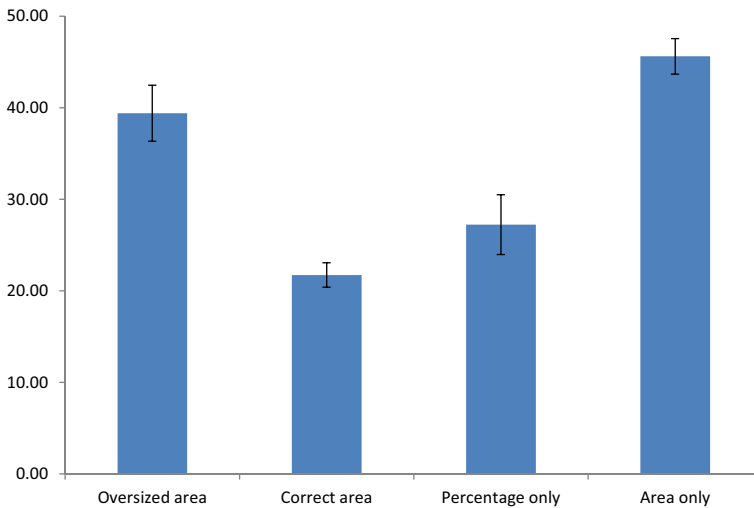
Table 2 presents the results of the ANOVA analyses on the measures of volume perception, as well as the underlying means per condition. Results indicated that the oversized area affected participants’ visual estimates of the bonus volume. Specifically, when estimating the additional volume visually, the estimates were larger for the package with an oversized area on the label than for the package with a correct area on the label (see also Fig. 1). This implies that the oversized area on the label increased the perceived additional volume that the package contained. Furthermore, although the difference in estimates between the oversized area and the area-only conditions did not reach statistical significance, there was a marginal difference between these conditions (at  $p = .076$ ). This suggested that participants were also somewhat affected by the percentage indicated on the package and adjusted their estimates downward to some extent when the area and the percentage were not aligned compared to when only the area was indicated. Together, this supported that notion that although participants used both anchors (the indicated percentage and the coloured area) in their estimates, the coloured area



**Table 2** Mean product perception in each of the conditions and ANOVA results

Dependent variable	Oversized area ( <i>n</i> = 29)	Correct area ( <i>n</i> = 33)	Percentage only ( <i>n</i> = 30)	Area only ( <i>n</i> = 34)	<i>F</i> -value ( <i>df</i> <sub>1</sub> = 3; <i>df</i> <sub>2</sub> = 122)	<i>p</i>	$\eta_p^2$
Visual estimate: Estimated area border (in mm from the top)	39.41 a (3.06)	21.73 b (1.33)	27.23 b (3.26)	45.62 a (1.94)	20.87	<.001	0.18
Perception of bonus volume (seven-point scale)	4.62 a (0.26)	4.06 a (0.18)	4.20 a (0.22)	5.32 b (0.18)	7.81	<.001	0.16
Estimated volume percentage (in %)	13.79 a (1.04)	13.97 a (1.63)	14.30 a (1.76)	22.47 b (1.31)	8.67	<.001	0.18
Perception of pack size (seven-point scale)	4.66 a (0.23)	4.91 a (0.21)	4.83 a (0.20)	5.07 a (0.17)	0.74	.533	0.02

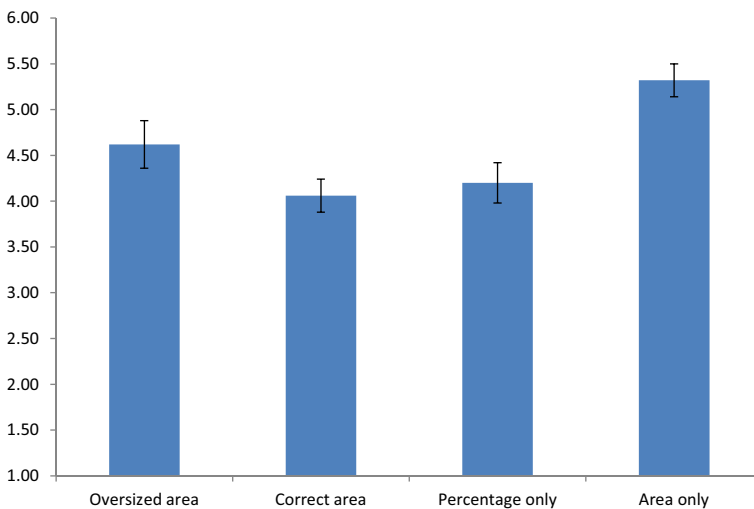
Number of participants in each condition between brackets in the table heading. Means with different letters significantly differ in the pairwise comparison test (LSD) at *p* < .05. Standard errors between brackets. Partial eta squared ( $\eta_p^2$ ) is a measure of the effect size



**Fig. 1** Effects on the estimated area border (in mm from the top)

was particularly influential and could thus deceive them into thinking that the extra volume was more than it actually was.

When participants indicated their perception of bonus volume on a seven-point scale, the results were also in line with our expectations (see Fig. 2 for a visual illustration). In this case, the difference between the oversized area and area-only conditions was significant, whereas the difference between the oversized area and correct area conditions was marginally significant (at  $p = .06$ ). In other words, both the area that was indicated and the percentage affected participants' estimates of the additional volume in the package. When presented with a pack



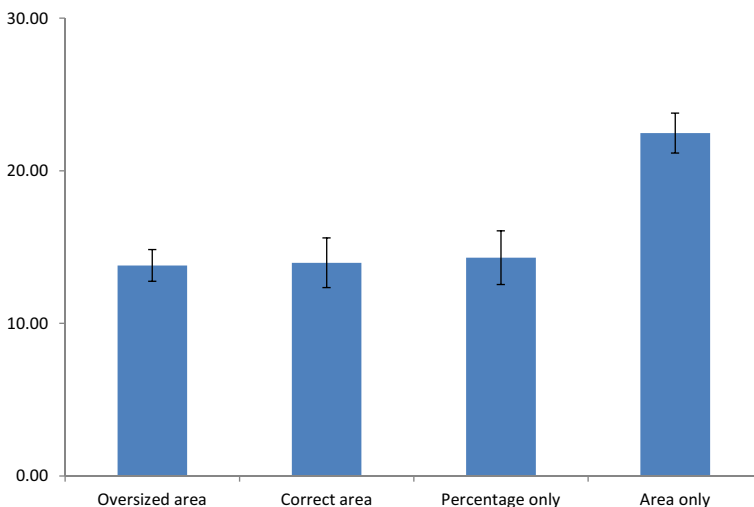
**Fig. 2** Effects on the perception of bonus volume

with an oversized area indication and a percentage, they perceived more bonus volume than in a package with the correct area indication, even though they perceived less bonus volume than in a package with only an oversized area.

Participants also indicated the *percentage* of bonus volume that (they thought) was in the pack. As expected, in the area-only condition, where no percentage was indicated on the package itself and the indicated area was large, participants reported a larger percentage than in the other conditions. More interesting is that we did not find significant differences between the other three conditions. Thus, participants appeared equally aware of the percentage in these conditions, even though they had indicated more bonus volume in the oversized area condition. It was also noteworthy that we did not find any differences between the condition in which the 10% area was indicated with a correct area (the correct area condition) and the condition in which only the percentage was shown without an area indication (the percentage-only condition). This indicates that participants were able to interpret the percentage. In other words, the results that we found did not appear due to a misinterpretation of the percentage number. A visual illustration of the results is provided in Figure 3.

Results furthermore showed that the perception of total pack size, which did not differ across conditions, was not significantly affected. Thus, participants did not perceive the total pack as larger when they were presented with an oversized area. Their estimates of which proportion of the pack was the bonus volume were affected, but not their perceptions of the total pack.

Next, we examined the area border that participants indicated in more detail. In the condition with an oversized area, 82.8% of participants indicated an area on pack that was larger than the objectively correct area (set at 20 mm from the top), and 51.7% overestimated by more than twice the size of the correct area (i.e., more than 40 mm from the top). This indicates that most people substantially overestimated the extra volume in the package. To further investigate this statistically, we performed *t* tests to examine whether the estimated area border (in mm from the top) differed from the actual area border for the bonus volume of 10%.



**Fig. 3** Effects on the estimated volume percentage

The estimates for the package with the oversized indication of bonus volume significantly differed from this latter number ( $t(28) = 6.35, p < .001$ ). Thus, also objectively, participants overestimated the area of bonus volume for this package. To examine if this overestimation occurred only for participants who did not remember the 10% indication on package, or also for participants who correctly reported the 10%, we split the data. Participants who did not correctly report the 10% number estimated the area border at 47.54 mm from the top, which was significantly different from the 20 mm ( $t(12) = 8.46, p < .001$ ). Participants who correctly reported the 10% number gave a lower estimated area border at 32.81 mm, which was still a significant overestimation of the actual bonus area ( $t(15) = 3.00, p < .01$ ). Thus, even participants who correctly believed that the bonus volume was 10% overestimated the area of this bonus volume.

Estimates in the other conditions were also compared to the objective number. The package in which only an area was indicated (larger than 10%) and no percentage was included on the package (the area-only condition) should lead to higher estimates than the 10% target, and results reflected this ( $t(33) = 13.24, p < .001$ ). Furthermore, estimates for the package in which the correct area was indicated did not significantly differ from the objectively correct number ( $t(32) = 1.30, p = .205$ ). Unexpectedly, estimates in the percentage-only condition were significantly different from the correct number ( $t(29) = 2.20, p < .05$ ). Thus, when only the percentage was indicated on the package, participants overestimated the area of bonus volume. This overestimation was smaller than in the oversized area condition and suggested that participants had some difficulty in translating the percentage to an area on pack.

## Discussion

In this study, we have tested whether consumer perceptions of a bonus package containing an oversized indication on the extra volume in the package (e.g., an indication of 10% extra on a coloured area on the package that is considerably larger than 10%) deceive the consumer. We have empirically analysed a situation that is similar to one in the ECJ's "Mars" judgment in the area of the average consumer. Results from the experiment, comparing a package with an oversized area indication to relevant control conditions, provide pertinent insights on how consumers in the "real world" as confined by the method used actually respond to such a package. This type of empirical information could be gathered for other cases and different circumstances as well. Our study can thus serve as an illustration of how an integration of behavioural analysis into the legal domain could work in the area of EU unfair commercial practices.

Our results indicate that the oversized area deceives consumers. When estimating the additional volume visually, consumers estimate that the additional volume is larger when the package contains an oversized area than when it contains the correct area on package. Moreover, they estimate this area as significantly larger than the objectively correct area. Even though a correct anchor was provided on the package (it contained the indication "10% free"), this did not stop consumers from also relying in part on the visually indicated area. Importantly, when the area on pack correctly indicated 10% of the actual package size, such an overestimation did not occur. This implies that the oversized area on the label increases the perceived additional volume that the package contains. Similar results as those obtained for the visual estimate of bonus volume are also obtained for the seven-point rating scale of the amount of extra bonus volume, showing that results do not depend on the measurement scale that is used.

This study did not test whether this deceiving practice also “causes or is likely to cause (... the consumer) to take a transactional decision” (see, e.g., Art. 6 (1) UCPD), which is an additional prerequisite to determine “misleading” in EU unfair commercial practices law. While other rules on unfair commercial practices law such as for example Art. 7 of the Food Information Regulation<sup>6</sup> do not state explicitly that “misleadingness” requires both deceptiveness and a causal link of this deception to the transactional decision of the consumer, Art. 6 (1) UCPD is very precise on this requirement. This legal prerequisite, wittingly or unwittingly, reflects insights from behavioural sciences that have shown that deviations from rationality can often be attributed to low opportunity costs in experimental settings (e.g., Smith and Walker 1993<sup>7</sup>). Hence, according to both, legal requirements and “good scientific practice” of consumer decision theory, a study such as this one needs to be succeeded by at least another one which tests whether the deceptiveness of the practice actually alters consumer buying behaviour. If it does, the outcome of these experiments can be used to inform the interpretation of the term “misleading” regardless of the fact whether consumers are misled on rational or irrational grounds.

## **Transferring These Findings to EU Unfair Commercial Practices Law: Anything to Learn?**

In this section, we deal with the question whether we can draw normative conclusions from these experimental findings for EU internal market regulation. We will first evaluate whether and if yes to what extent one can draw normative conclusions from the specific settings of a behavioural study such as the one at hand for EU internal market regulation.

Whether these findings can be used to solve questions of regulation depends to a large extent on whether they are generalizable (for a discussion of the issue of generalizability in experimental consumer research, see, e.g., Lynch 1982) and if they fit into the context of the regulatory regime at play, the so-called legal validity test (see Purnhagen and Feindt 2015). It is in the nature of regulation that it requires the lawmaker or regulator to find a standardized protective level that applies for all addressees (Sunstein 1997). It is likewise in the nature of empirical research that any single study, this one included, should be evaluated with care to estimate the likely generalizability of findings to other consumers, countries or products. The issue of generalizability relates to participants, stimulus material (product) and measurement method, while the question of legal validity relates to underlying value commitments in the law, both of which we will discuss in turn.

### **1. Generalizability**

First, participants in our study consist of students, which implies that education level is higher than that of the general population. Moreover, participants were given ample opportunity to evaluate the product without distraction. Students, under these circumstances, may be better able or more likely to use the verbal percentage anchor than people with a lower education level or with other demands on their attention. After all, there are large individual differences in the extent to which people have the numeracy skills to

<sup>6</sup> Regulation (EU) No. 1169/of the European Parliament and the Council of 25 October 2011 on the provision of food information to consumers, OJ L 304, 22.11.2011, p. 18–63.

<sup>7</sup> We would like to thank the anonymous reviewer for pointing this out.

correctly interpret percentages (Dowker 2005). This potentially diminishes the extent to which an overestimation of the bonus volume occurs in our sample compared to the general population, and effects could potentially be stronger in the general population, as future research could examine.

Second, this study presents a single empirical study in one European country. The EU is characterized by a rich and diverse European culture, which, according to Art. 3 (3) subpara 4 Treaty on European Union (TEU), needs to be taken into account when regulating the internal market (Smits 2007; Helleringer and Purnhagen 2014). In order to form a sound basis for regulatory decisions, further replication studies in other countries would need to follow. Yet, it is expected that these results would not differ significantly to the current one as our results are in line with a general principle of anchoring and adjustment, which has been demonstrated repeatedly across cultures and applications.

Furthermore, this study has examined one package of coffee, with a specific percentage and area indicated on the package. Although this is a common product, and both percentage and area are in line with the prominent Mars case, generalizations to other products and other indications of percentages and areas should be taken with care. Specifically, although anchoring effects have been demonstrated with unrealistically high anchors (in this case, for instance, an indicated area of, say, 70% of the package), results might be attenuated when such an anchor is used (Furnham and Boo 2011, pp. 35–42). If a court were to encounter cases in which the percentage and/or area is considerably different from the current example, additional empirical testing would be advisable as input for the court ruling.

Finally, with respect to generalizability across measurement methods, the current study provides positive evidence. Three measures of volume perceptions were included and results provided robust and consistent insights across these measurements.

A further point for consideration is that participants have indicated their perceptions of additional volume immediately after viewing the single package when the package itself was no longer in view. The study therefore cannot provide an indication of whether the effects occur to the same extent when packages are encountered in the context of competing products (as in a supermarket) or whether consumer responses are the same when they can view the package while providing their estimates. These questions are left for future research, although the robustness of anchoring effects at least suggests that results might not differ substantially. The advantage of the empirical approach is that such questions surface and can be addressed, in order to gain more insight over time.

## 2. Legal validity

The average consumer benchmark has been introduced as an operationalized test of the proportionality principle, weighing freedom of goods, Member State cultures, and consumer protection (Franck and Purnhagen 2014). It is hence, as stipulated in Recital 18 UCPD, not a statistical test but rather a normative benchmark (Franck and Purnhagen 2014). Whether this precludes the European judiciary to take into account findings from behavioural science or at least requires national Courts to take into account findings from behavioural science when interpreting the consumer benchmark is subject to debate (see more elaboration in Purnhagen 2017). The indicators that Recital 18 UCPD does not preclude such an incorporation of behavioural sciences are increasing: First and foremost, by illustrating only what can be “reasonably” expected

from an average consumer, the Court has introduced a far less rigorous notion of normativity with regard to consumer processing of information than regularly claimed (Franck and Purnhagen 2014, p. 338). Second, in the labelling case *Bundesverband der Verbraucherzentralen und Verbraucherverbände—Verbraucherzentrale Bundesverband e.V. v Teekanne GmbH & Co. KG*,<sup>8</sup> the CJEU interpreted the average consumer benchmark in the Food Information Regulation not in line with previous case law, but rather in line with findings from behavioural science (Schebesta and Purnhagen 2016, p. 595). Third, in *Scotch Whisky Association*,<sup>9</sup> the CJEU lately required from national courts to ex-officio apply newest scientific evidence in the suitability review of the proportionality test (for a comprehensive analysis to this end, see Alemanno 2016, pp. 1037–1064; Purnhagen and Schebesta 2017). If the proportionality principle gets “topped up” with scientific elements, this might have implications on the interpretation of the average consumer benchmark, which is by itself a product of the proportionality test (Purnhagen 2017). While it is hence far from clear whether and at what level of the current consumer benchmark requires or demands the incorporation of behavioural sciences according to EU law, it does not automatically preclude it.

## Conclusion

In situations similar to the Mars case, consumers are systematically deceived due to the anchoring bias. Such insights from behavioural findings can provide important information for Courts and regulators when interpreting the consumer benchmark to be applied in unfair commercial practice law. While consumer protection levels *in concreto* are still to be determined at Member State level, one could argue that systematic deviations could be dealt with at Union level, keeping the insights of behavioural sciences in mind. That is to say, regulators and Courts need to base their interpretation or exercise of discretion with an eye on multiple factors, of which the consumer benchmark as determined by behavioural science is one. From the empirical study, we can tentatively draw the conclusion that the anchoring bias for regulation appears to be a robust tool to determine whether consumers are systematically deceived. When bringing the vulnerable consumer and the classical concept of the average consumer in line, we hence argue that this indicates that Union regulators can rely on the anchoring bias in marketing law to determine the consumer benchmark. Whether this will lead Courts and regulators to adjust their decisions depends on the weight that regulators will give to insights from behavioural research to protect consumer versus other factors such as the provision of the principle of free movement of goods, when interpreting the law or exercising the margin of discretion that they enjoy when conducting the respective regulatory decisions.

Summarizing, insights from behavioural studies can inform regulators about the effects of commercial practices on consumer perceptions, and given the changes in the internal market concept, these insights are relevant for them to consider.

<sup>8</sup> Case C–195/14 *Bundesverband der Verbraucherzentralen und Verbraucherverbände—Verbraucherzentrale Bundesverband e.V. v Teekanne GmbH & Co. KG*, nyr.

<sup>9</sup> Case C–333/14 *Scotch Whisky Association and Others v The Lord Advocate and The Advocate General for Scotland*, nyr, para 56 f.

## Appendix. Packages used in the experiment

<p>Oversized area. Text in yellow area: “Temporary: +10% FREE”</p>  <p>The image shows a coffee package with a large yellow banner at the top. The banner contains the text "TIJDELIJK +10% GRATIS" in red. Below the banner, the package features the brand name "KANIS GUNNINK" in large black letters, a circular logo with "KG" and "SINCE 1885", and the text "CAFÉÏNEVRIJ" and "SNELFILTERMALING" at the bottom.</p>	<p>Correct area. Text in yellow area: “Temporary: +10% FREE”</p>  <p>The image shows a coffee package with a yellow banner at the top. The banner contains the text "TIJDELIJK +10% GRATIS" in red. Below the banner, the package features the brand name "KANIS GUNNINK" in large black letters, a circular logo with "KG" and "SINCE 1885", and the text "CAFÉÏNEVRIJ" and "SNELFILTERMALING" at the bottom.</p>
<p>Percentage only. Text in yellow circle: “+10% FREE”</p>  <p>The image shows a coffee package with a yellow circle on the top left. The circle contains the text "+10% FREE" in black. Below the circle, the package features the brand name "KANIS GUNNINK" in large black letters, a circular logo with "KG" and "SINCE 1885", and the text "CAFÉÏNEVRIJ" and "SNELFILTERMALING" at the bottom.</p>	<p>Area only. Text in yellow area: “Temporary: Extra Contents”</p>  <p>The image shows a coffee package with a yellow banner at the top. The banner contains the text "TIJDELIJK Extra inhoud" in red. Below the banner, the package features the brand name "KANIS GUNNINK" in large black letters, a circular logo with "KG" and "SINCE 1885", and the text "CAFÉÏNEVRIJ" and "SNELFILTERMALING" at the bottom.</p>



**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Alemanno, A. (2016). Balancing free movement and public health: The case of minimum unit pricing of alcohol in Scotch Whisky. *Common Market Law Review*, 53, 1037–1064.
- Amir, O., & Lobel, O. (2012). Liberalism and lifestyle: Informing regulatory governance with behavioural research. *European Journal of Risk Regulation*, 3, 17–25.
- Burgess, A. (2012). “Nudging” healthy lifestyles: The UK experiments with the behavioural alternative to regulation and the market. *European Journal of Risk Regulation*, 3, 3–16.
- Chandon, P., & Ordabayeva, N. (2009). Supersize in one dimension, downsize in three dimensions: Effects of spatial dimensionality on size perceptions and preferences. *Journal of Marketing Research*, 46, 739–753.
- Chen, H. A., & Rao, A. R. (2007). When two plus two is not equal to four: Errors in processing multiple percentage changes. *Journal of Consumer Research*, 34, 327–340.
- Chen, H. A., Marmorstein, H., Tsiros, M., & Rao, A. R. (2012). When more is less: The impact of base value neglect on consumer preferences for bonus packs over price discounts. *Journal of Marketing*, 76, 64–77.
- Diamond, W.D. (1992). Just what is a “dollar’s worth”? Consumer reactions to price discounts vs. extra product promotions. *Journal of Retailing*, 68, 254–270.
- Dowker, A. (2005). *Individual differences in arithmetic: Implications for psychology, neuroscience and education*. Hove and New York: Psychology Press.
- Duivendoorde, B. (2015). *The consumer benchmarks in the unfair commercial practices directive*. Cham et al.: Springer Science.
- Englich, B. (2006). Blind or biased? Justitia’s susceptibility to anchoring effects in the courtroom based on given numerical representations. *Law & Policy*, 28, 497–514.
- Franck, J.-U., & Purnhagen, K. (2014). Homo economicus, behavioural sciences, and economic regulation: On the concept of man in internal market regulation and its normative basis. In K. Mathis (Ed.), *Law and economics in Europe—Foundations and applications* (pp. 329–365). Dordrecht: Springer Science.
- Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics*, 40, 35–42.
- Hardesty, D. M., & Bearden, W. O. (2003). Consumer evaluations of different promotion types and price presentations: The moderating role of promotional benefit level. *Journal of Retailing*, 79, 17–25.
- Heise, M. (1999). The importance of being empirical. *Pepperdine Law Review*, 26, 808–834.
- Helleringer, G., & Purnhagen, K. (2014). On the terms, relevance and impact of a European legal culture. In G. Helleringer & K. Purnhagen (Eds.), *Towards a European Legal Culture* (pp. 3–14). Munich, Oxford, Baden-Baden: C. H. Beck, Hart, Nomos.
- LeBoeuf, R. A., & Shafir, E. (2006). The long and short of it: Physical anchoring effects. *Journal of Behavioral Decision Making*, 19, 393–406.
- Lynch, J. G. (1982). On the external validity of experiments in consumer research. *Journal of Consumer Research*, 9, 225–239.
- Mathis, K. (Ed.) (2015). *European perspectives on behavioural law and economics*. Cham et al.: Springer Science.
- Mishra, A., & Mishra, H. (2011). The influence of price discount versus bonus pack on the preference for virtue and vice foods. *Journal of Marketing Research*, 48, 196–206.
- Mussweiler, T., Strack, F., & Pfeiffer, T. (2000). Overcoming the inevitable anchoring effect: considering the opposite compensates for selective accessibility. *Personality and Social Psychology Bulletin*, 26, 1142–1150.
- Oppenheimer, D. L., LeBoeuf, R. A., & Brewer, N. T. (2008). Anchors aweigh: A demonstration of cross-modality anchoring and magnitude priming. *Cognition*, 106, 13–26.
- Price, D. D., Bush, M., Long, S., & Harkins, S. W. (1994). A comparison of pain measurement characteristics of mechanical visual analogue and simple numerical rating scales. *Pain*, 56, 217–226.
- Purnhagen, K. (2015). Why do we need responsive regulation and behavioural research in EU internal market law? In K. Mathis (Ed.), *European perspectives on behavioural law and economics* (pp. 51–69). Cham et al.: Springer Science.
- Purnhagen, K. (2017). More reality in the CJEU’s interpretation of the average consumer benchmark—also more behavioural science in unfair commercial practices? *European Journal of Risk Regulation*, 8.

- Purnhagen, K., & Feindt, P. (2015). Making behavioural insights work for the commission's new better regulation strategy. *European Journal of Risk Regulation*, 6, 361–368.
- Purnhagen, K., & Schebesta, H. (2017). A case moving at the Frontiers of market access, freedom of goods, the common agricultural policy and science in court-reflections on Scotch Whisky Association. *European Law Review*, 54.
- Robbenolt, J. K., & Studebaker, C. A. (1999). Anchoring in the courtroom: The effects of caps on punitive damages. *Law and Human Behavior*, 23, 353–373.
- Schebesta, H., & Purnhagen, K. (2016). The behaviour of the average consumer: A little less normativity and a little more reality in Court's case law? *Reflections on Teekanne*, *European Law Review*, 14, 590–598.
- Schuck, P. H. (1989). Why don't law professors do more empirical research? *Journal of Legal Education*, 39, 323–336.
- Sibony, A.-L. (2015). Can EU consumer law benefit from behavioural insights?: An analysis of the Unfair Practices Directive. In K. Mathis (Ed.), *European perspectives on behavioural law and economics* (pp. 71–106). Cham et al.: Springer Science.
- Sibony, A.-L., & Alemanno, A. (2015). The emergence of Behavioural policy-making: A European perspective. In A.-L. Sibony & A. Alemanno (Eds.), *Nudge and the law—a European perspective* (pp. 1–25). Oxford: Hart.
- Smith, V.L., & Walker, J.M. (1993). Monetary rewards and decision cost in experimental economics. *Economic Inquiry*, pp. 245–261.
- Smits, J. (2007). Law making in the European Union: On globalization and contract law in divergent legal cultures. *Lousiana Law Review*, 67, 1181–1203.
- Stuyck, J. (2010). Setting the scene, chapter I. In H. Micklitz, J. Stuyck, E. Terry, & D. Droshout (Eds.), *Consumer law: Ius commune casebooks for a common law of Europe* (pp. 1–68). Oxford: Hart.
- Sunstein, C. (1997). *Free market and social justice*. Cambridge: Cambridge University Press.
- Thaler, R., & Sunstein, C. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven: Yale University Press.
- Trzaskowski, J. (2011). Behavioural economics, neuroscience, and the unfair commercial practices directive. *Journal of Consumer Policy*, 34, 377–392.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124–1131.
- Van Aaken, A. (2015). Judge the nudge: In search of the legal limits of paternalistic nudging in the EU. In A.-L. Sibony & A. Alemanno (Eds.), *Nudge and the law—A European perspective* (pp. 83–112). Oxford: Hart.
- van Bavel, R., Herrmann, B., Esposito, G., & Proestakis, A. (2013). *Applying behavioural sciences to EU policy-making*, JRC Scientific and Policy Reports. Seville: Joint Research Center European Commission, available at [http://ec.europa.eu/dgs/health\\_consumer/information\\_sources/docs/30092013\\_jrc\\_scientific\\_policy\\_report\\_en.pdf](http://ec.europa.eu/dgs/health_consumer/information_sources/docs/30092013_jrc_scientific_policy_report_en.pdf).
- Vanhuele, M., & Drèze, X. (2002). Measuring the price knowledge shoppers bring to the store. *Journal of Marketing*, 66, 72–85.
- Wansink, B., & van Ittersum, K. (2003). Bottoms up! The influence of elongation on pouring and consumption volume. *Journal of Consumer Research*, 30, 455–463.
- Weatherill, S. (2007). Who is the 'average consumer'? In S. Weatherill & U. Bernitz (Eds.), *The regulation of unfair commercial Practices under EC Directive 2005/29* (pp. 115–138). Oxford: Hart.
- Whyte, G., & Sebenius, J. K. (1997). The effect of multiple anchors on anchoring in individual and group judgment. *Organizational Behavior and Human Decision Processes*, 69, 75–85.
- Xu, Y., & Huang, J.-S. (2014). Effects of price discounts and bonus packs on online impulse buying. *Social Behavior and Personality*, 42, 1293–1302.
- Zamir, E., & Teichman, D. (Eds.). (2014). *The Oxford handbook of behavioral economics and the law*. Oxford: Oxford University Press.