# Measuring the Price Knowledge Shoppers Bring to the Store

Marc Vanhuele
Xavier Drèze*

January 2001

*Marc Vanhuele is Associate Professor of Marketing at HEC Paris (France), and Xavier Drèze is Visiting Professor of Marketing at the University of California, Los Angeles. The authors thank the HEC Foundation for its financial support, and Shantanu Dutta, Gilles Laurent, four anonymous reviewers, and the JM editor for their constructive feedback.

## Abstract

Reference price research suggests that consumers memorize and recall price information when selecting brands for frequently purchased products. Previous price-knowledge surveys, however, indicate that memory for prices is lower than expected. In this study, we show that these price-knowledge surveys actually provided imperfect estimates of price knowledge because they focused only on recall and short-term memory. We propose, instead, to use a combination of *price recall*, *price recognition*, and *deal recognition* to measure the degree to which consumers use *auditory verbal*, *visual Arabic*, or *analogue magnitude* representations to memorize prices. We show how the combination of these three measures provides a much richer understanding of consumer's knowledge of prices. Our results suggest that the price knowledge involved in reference prices may often not be accessible to recall but shows up in price recognition and deal recognition. In addition we identify consumer and product characteristics that explain the variations in price knowledge. We find, for instance, that frequent promotions increase the ability of consumers to remember regular prices and that store switchers do not possess a better price knowledge than other shoppers.

Consumers have a strong interest in keeping a knowledge base of prices for products they frequently purchase. This knowledge base permits them to assess the attractiveness of advertised promotions (in flyers, advertisements, and the store itself), it alerts them to price increases, and it allows them to compare prices across stores. Marketers are evidently interested in finding out how complete and accurate this knowledge base is. The assumption of early economic price theory that consumers are aware of most prices has clearly been invalidated by surveys of price knowledge. The in-store price knowledge surveys by Dickson and Sawyer (1990) provided estimates of price knowledge that were surprisingly low. Only 47 to 55% of the respondents could accurately recall the price of an item they had just placed in their shopping cart and 19 to 23% did not even attempt to give an estimate. These results were replicated by other researchers (Le Boutillier, Le Boutillier, and Neslin 1994; Wakefield and Inman 1993).

A consumer's price knowledge base is, by definition, stored in *long-term memory*. Although the results of the Dickson and Sawyer type of studies are often interpreted as estimates of consumer price knowledge, their interview procedure actually targets in-store attention to prices, by checking the contents of *short-term memory*. The power of their demonstration is based on the fact that, because price recall questions are asked immediately after the consumer picked a product from the shelf, "it seems unlikely that shoppers noted the price and forgot it in the short intervening time" (p. 50). It is obvious that, although a respondent may call on long-term memory to give a price estimate, short-term memory will clearly dominate the results. This implies that Dickson and Sawyer may considerably *overestimate* the extent to which consumers have long-term knowledge of prices of frequently purchased products. Interestingly, however, it is also possible that the authors *underestimate* price knowledge, because they rely on price recall. As explained by Monroe and Lee (1999), memory for price information may not be recallable and

still be useful in making price judgments. Instead of relying on recall, consumers may be alerted to price changes by recognition (or more precisely the lack of recognition) and recognition performance is usually better than recall performance. Moreover, Monroe and Lee also argue that price knowledge in memory may not be accessible to recall *and* recognition and still be instrumental in shopping because it gives a sense of magnitude of what the normal price is. In fact, memory research has focused in the past decade on this form of memory (referred to as implicit memory).

To measure the long-term price knowledge which consumers bring to the store clearly requires a different interview methodology than that used by Dickson and Sawyer. The main objective of our project is therefore to design and test a methodology that taps into the different forms of long-term memory for prices. We base our design on the recent literature in the area of numerical cognition that shows that numerical information can be represented in three different forms in memory. Only one of these forms is directly accessible to verbal recall. Getting a complete picture of price knowledge therefore requires the use of measures that tap into all three forms of representation. In a survey of 400 consumers in a French supermarket, we measure price knowledge with this procedure and, in addition, examine to what extent price knowledge is driven by the characteristics of product categories and consumer shopping strategies. Diagnosing levels of price knowledge and differences across consumers and across product categories is important for managers who have to coordinate prices in a competitive multi-category environment. Academics will, in addition, be interested in the relationship between the different measures of long-term memory for price and in the determinants of price knowledge in a low-involvement purchase context. They will also be particularly interested in the implications of our

research for the apparent conflict between the results of reference price studies and those of price knowledge surveys.

Reference price models are choice models, estimated on scanner data, that include the effect of price knowledge on product choice through a surrogate measure (Winer 1986; Kalyanaram and Winer 1995). In the absence of direct measures of price knowledge, these models measure the reference price of a certain consumer for a certain product as a combination of the past actual prices of this product at previous shopping occasions. The underlying assumption is that the consumer was exposed to the product's price at these shopping occasions, and therefore may have memorized it. This assumption is, however, clearly put into question by past price knowledge surveys. Several authors have therefore raised the question of how reference price models can "significantly predict brand choice if actual market prices are often not noticed or remembered by consumers" (Urbany and Dickson 1991, p. 51)? Price survey results have, however, also been interpreted as an indication that "a sizeable proportion of consumers are capable of recalling prices" (Briesch et al. 1997; also see Kalyanaram and Winer 1995). In this paper we argue that, in fact, neither interpretation is satisfactory because the Dickson and Sawyer type of price knowledge surveys focus on short-term memory for price while reference price is, by definition, stored in long-term memory. Moreover, as explained in the previous paragraph, reference price does not have to be recallable to exert its effect on choice. By examining the different forms of long-term memory for price through multiple measures, our study helps in advancing the debate between reference price and price knowledge survey research. We show that recall-based surveys severely underestimate the level of price knowledge available in the market. Our study also indicates that reference prices may be based on forms of knowledge that are not easily recallable.

We start this paper by examining, from a theoretical perspective, how consumers store prices in memory and, consequently, which measures should be used to determine the level of price knowledge. In the second section of the paper, these measures are included in a survey procedure that measures the different dimensions of long-term consumer price knowledge that consumers bring to the store. The third section introduces the factors that drive price knowledge and the fourth section explains the methodology we used in detail. We then present the results of the survey we ran with 400 respondents in a French supermarket and discuss these results in the final section of the paper.

## Long Term Memory and Number Representation

To develop measures of price knowledge, we first must understand how prices are stored in long-term memory. Research on numerical cognition focuses on how numbers and arithmetic information are represented in memory and used in cognitive arithmetic tasks (Ashcraft, 1992; McCloskey and Macaruso, 1995). As a synthesis of the essential findings of this research, Dehaene (1992; also see Dehaene and Akhavein, 1995) develops a triple-code model where numbers can be mentally represented and manipulated in three different forms. The auditory verbal code manipulates a word sequence (e.g., /thirty/ /five/), the visual Arabic code represents numbers on a spatial visual medium (e.g., 35), and the analogue magnitude code represents numbers as approximate quantities on a dimension termed the number line (about 35, or somewhere between 30 and 40).

The triple-code model implies that prices can have three types of representation in memory. The auditory verbal code, which is for instance the code involved in verbal counting and arithmetical fact retrieval, can be used to store and retrieve price information that has been —often sub-

vocally— read out by the consumer. Alternatively, the same information can also be stored as a string of visual symbols. Comparing two numbers, for instance, usually involves the visual Arabic code. There is evidence that some individuals are auditory and others visual types, but one individual may store some prices in auditory and others in visual form, depending on how the information was coded initially. The same price can also be stored in multiple forms. Which code is accessed at a certain point in time depends on the task at hand. Although, obviously, conversions can be made from one code to the other, there is a cognitive cost involved in this translation, which is revealed in increasing reaction times. The analogue magnitude code, which is the third type of representation of price in memory, is somewhat different from the other two codes because it is formed on the basis of auditory or visual input (i.e. the information passes through the other two codes first), through a fairly automatic process that recodes a number into an approximate quantity.

Important for our study is that Dehaene posits that each number-related task is tied to a specific code. A verbal recall question is most readily answered through accessing the auditory verbal code. For a recognition question, in which a price is shown visually, accessing the visual Arabic code is the most evident process, while estimates of the normal price range or judgments of the attractiveness of a certain price can be based on the magnitude code. As explained in the previous paragraph, transcoding from one code to another is possible, but requires mental switching between notations and increases response time. In face of a certain numerical task, the most evident code will therefore be privileged.

By using a combination of three price knowledge questions, one for each code, we can take the three possible forms of representation of price knowledge into account and get a complete and accurate picture of the level and types of price knowledge a certain consumer holds. Although the

literature on numerical cognition was the source of inspiration for the design of our survey procedure, we have to note that the three ways of measuring price knowledge have also been discussed individually in the marketing literature. Recall has been used most often in past studies, but Monroe et al. (1986) have argued that recognition is a more appropriate measure of price knowledge than recall. They explain that price information can be the result of active search or it can be learned incidentally, almost by chance, while shopping. Active search may be more the exception than the rule, but it is more likely to lead to recallable price information, because it makes explicit links to information already stored in memory. Just picking a product from the shelf and incidentally noticing its price, in contrast, may not lead to recall of this price, even after several repetitions, because no such links to memory are made. Memory researchers refer to these forms of price information processing as inter-item processing (several prices are compared) and intra-item processing (one price for one product is examined) (Mandler 1980). If incidental learning is more prevalent, recognition may indeed be more appropriate than recall as a measure of price knowledge.

Regarding the notion that looking for price magnitudes instead of precise price points may in some cases be more appropriate, there is research suggesting the existence of reference price regions instead of price points (e.g., Kalyanaram and Little 1994). Small differences in reference price do not seem to have any effect in some choice models. In addition, Monroe and Lee (1999) explain how the notion of implicit memory expresses the ability to judge the attractiveness of a certain price, without the possibility of explicit recall of an actual price.

## A Survey Procedure to Measure Price Knowledge

Based on our literature review, we propose that an investigation of long-term price knowledge needs to tap into all three memory codes for numerical knowledge. In our survey of consumers we measure three constructs that correspond to these three forms of price knowledge:

- *Recallable price knowledge*: the consumer knows the actual price of the product in the store "by heart" (Urbany and Dickson 1991). This is the highest level possible and is mainly based on the *auditory verbal* code.
- *Price recognition*: unaided price knowledge is not present, but aided price recognition is (Monroe et al. 1986). When seeing a price on a product, people can tell if this is the price they are used to and have in mind, by accessing the *visual Arabic* code. If no promotional signal is present, a lack of recognition usually implies that the price has been increased.
- *Deal spotting*: this form of price knowledge consists of noticing that a price is within or outside the normal range of previous prices, for which accessing the *analogue magnitude* code is sufficient (Monroe and Lee 1999). People with this level of knowledge do not really know by heart what products cost, and they cannot tell whether a presented price is exactly the one they are used to. They can, however, recognize a good deal or a bad deal when they see one. They have, in other words, a sense of what the normal price range is, if presented with sufficient cues.

Because our objective is to measure the price knowledge consumers bring to the store and because we want to draw implications for reference price research, our survey procedure has to match as closely as possible the context of an actual in-store reference-price comparison. In addition to the three-step memory questions, our survey procedure has the following key features:

- Our questions are *specific to the stock-keeping unit (SKU)* purchased and are for SKUs that have been purchased in the past by the respondents (unlike, for instance, Urbany and Dickson 1991).
- The questions tap into *long-term* memory and exclude the possibility that price is retrieved from short-term memory (unlike Dickson and Sawyer 1990 and their replications).
- We match *product and brand cues* used in the questioning to those present in the choice context (e.g., through the use of photographs of the products tested).
- The questioning takes place in the store, at the start of a purchase occasion in order to maximize the number of *contextual cues* and to ensure the presence of normally available shopping knowledge (as opposed to in-home surveys like that of Urbany and Dickson 1991 which may underestimate actionable price knowledge).

In addition to our main survey of long-term price knowledge, we ran an additional survey that measured consumers' memory for prices immediately after they picked a product from the shelf. This survey replicates the work of Dickson and Sawyer and is meant to benchmark our results to theirs in order to elucidate the debate we alluded to in the introduction between researchers who worked on scanner data with reference price models and those who used surveys.

## Drivers of Price Knowledge Accuracy

### *Category-Level Drivers of Price Knowledge*

One can readily see why a category's characteristics might influence the ability of a consumer to accurately store and use the price knowledge of one of its products. If a category had only one item, and if the price of that item never changed, it would not be very difficult for an individual who purchases from that category on a regular basis to remember the price of the item. In

contrast, if the category had dozens of different items and very volatile prices, remembering the price of any given product might be more difficult. We focus here on three dimensions of a category that might affect the consumer's ability to process price information accurately: price volatility, price range, and category clutter. Price volatility measures how often prices change over time. It is a function of the level of promotional activity in the category. Price range is measured as the difference in price from the highest to the lowest priced SKU in the category. Category clutter is operationalized as the number of SKUs in the category. Underlying our hypotheses for the effect of these factors is the notion that increased complexity in price information has a negative impact on memory performance.

When a product is often on promotion, its promoted prices will become more salient, and recall and recognition of the normal price should be affected negatively (Johnson, 1994). However, if a product is often promoted on price, its normal price range should be salient (Kalyanaram and Little, 1994) and deal recognition, which involves reacting to prices differences that exceed the typical price promotions (e.g., -20%), should therefore be enhanced. Therefore, our prediction for price volatility depends on the measure considered. Price recall ($H_{1a}$) and price recognition ($H_{1b}$) of the normal price should be lower in often promoted product categories while deal spotting should be higher ($H_{1c}$) (see appendix A for the formal hypotheses). When the price range is large, the product category will be associated in memory with a large number of different prices, and the category therefore becomes an unreliable cue to access any given price in memory (Anderson and Bjork 1994). This should lead to a higher level of confusion about the actual price of any given brand. Thus, prices in categories with a large price range should be more difficult to remember ($H_{2a}$-$H_{2c}$), be it through active or incidental learning. Category clutter is hypothesized

to have a similar effect, because confusion again becomes more likely when the number of SKUs in a category increases ($H_{3a}$-$H_{3c}$).

***Consumer-Level Drivers of Price Knowledge***

Previous studies have examined some of the factors that explain the differences among consumers in their price search (e.g., Urbany, Dickson, and Kalapurakal 1996) and price knowledge (Dickson and Sawyer 1990 and their replications). To complement this work, we divided consumer characteristics into three groups of variables: the propensity to engage in in-store price search, the propensity to engage in across-store price search, and shopping trip/household size. For each variable we developed a set of survey items.

Consumers who extensively engage themselves in price search should obviously also be the ones having the best knowledge of prices. They apparently use price as a decision variable (be it to compare brands within a store or prices across stores), and they also process more price information than consumers who do not engage in price comparisons. Therefore, we formulate the hypothesis that consumers will have more price knowledge when they engage in in-store price search ($H_{4a}$-$H_{4c}$) or across-store price search ($H_{5a}$-$H_{5c}$) than when they do not (Dickson and Sawyer 1990).

In terms of average shopping-trip size (at the household level), it is obvious that there is more to be gained from accurate price knowledge when the shopping basket is large than when it is small. Hence we can hypothesize that consumers engaging in larger average shopping trips will have better price knowledge than those engaging in small average trips ($H_{6a}$-$H_{6c}$).

## Price Knowledge Survey

### *Product Category Selection*

The product categories for the survey were chosen so that they represent high or low (but not medium) levels of each of the three factors we studied (price volatility, price range and category clutter). We selected a representative product category for each cell of our 2x2x2 factorial design in two steps. A first selection of possible candidates was made based on estimates from a national store panel of about three hundred product categories. We selected categories that were in either the first quartile (low) or the fourth quartile (high) in terms of price volatility over time, average price range, and number of SKUs. After this first selection we ran three store checks over a period of 2 months in the store where our survey took place to determine the eight most representative product categories (see Figure 1).

Each respondent had to answer price knowledge questions for three of the eight products. The interviewers had a list that determined the selection and order of the products for each interview. This way all products were presented approximately the same number of times. In case a respondent did not normally buy a given product in a super or hypermarket, the next one on the list was examined.

### *Respondent Selection and Interview Procedure*

The interviews all took place in one hypermarket. It was selected because its prices were almost exactly at the mean of 1,524 super and hypermarkets that were compared in a store price-check by a leading French consumer organization (Que Choisir 1998). The objective of our respondent selection was to interview a representative sample of regular shoppers of this store. Respondents were intercepted at one of the store entrances. As soon as an interviewer became available he or she had to solicit the third person that entered the store for participation in the interview. In total

400 shoppers were interviewed. Because we expected to find different types of shoppers at different times of the day and the week, the interviews were scheduled such that we covered each relevant time slot (morning, midday, evening; beginning of the week, normal weekday, and weekend).

The interview was described as being part of a study on consumer products and the prospective respondents were told they would receive a store coupon of 20FF at the end of the interview[1]. To qualify for the interview, shoppers had to pass two filter questions: they had to normally do their shopping themselves, and they had to do their regular shopping in the store of the interview. We assumed that people who were not regular shoppers would not have reliable reference prices. The problem with those who do not do their regular shopping at the store of the interview is that we have no practical way of knowing what the actual prices are at the store they patronize on a regular basis (i.e., they might have a very accurate price knowledge, but for prices that are different from the ones in effect at our store).

As already mentioned, our main survey was designed to assess price knowledge at a shopping occasion before contact with product prices. This was done to keep respondents from accessing data from short-term memory while still providing them with environmental cues to help retrieval from long-term memory. Pictures were taken of the 174 SKUs sold by the store in the eight product selected categories. These were reproduced in an interview folder sorted in alphabetical order by brand, and, at a second level, by SKU. Respondents were only interviewed on product categories they usually buy at the store. After a respondent had indicated in the interview folder which SKU s/he usually purchases, s/he had to give an estimate of its normal non-promoted price. Respondents could give the price they recalled in the unit they preferred: by item, by pack, by weight, or by volume. Next, a series of hypothetical prices for the product were presented in

the interview folder, one at a time. The respondent had to indicate whether each price represented a good deal, a normal price, or a bad deal. The price presented in this deal-spotting question was a unit price accompanied, where relevant (e.g., for mineral water) with a price per weight or volume. Respondents were assigned to one of two conditions for the deal-spotting questions. In the first condition, the price series started with a price that was 20% below the actual price, and followed with prices that were 5% below, 5% above, and 20% above the actual price. In the second condition, the price series started with a price that was 20% higher than the actual price, and followed with prices that are 5% above, 5% below, and 20% below the actual price. The two conditions were counterbalanced across the different interview folders such that each product category appeared in each series approximately the same number of times. The selection of the product categories and the counterbalancing of the price presentations was arranged such that respondents received different sequences of prices in the deal spotting questions and could therefore not predict whether the four consecutively presented prices would be decreasing or increasing.

In the following step the actual price of the product was shown, along with a 10% higher price and a 10% lower price. The interviewer told the respondent that one of the three prices was the normal price at the store and asked to indicate which one it was. The presentation order of the three prices in this recognition question was counterbalanced across interview folders.
The series of questions was repeated for two other product categories and the price knowledge questions were followed by twelve questions about the shopping habits and identity of the respondent. To maximize comprehension of the task, all price presentations, questions, and possible answers were presented orally, as well as in writing.

## Results

In this section we proceed in two steps. First we report the overall results for each of the three measures of price knowledge of French consumers. Next, we examine the degree to which these measures represent incremental levels of price knowledge. Second, we investigate the extent to which these results are category dependent and driven by the demographic characteristics and shopping habits of our respondents.

### *Measures of Price Knowledge*

#### Price Recall

Table 1 shows the percentage of responses that matched the actual price of the products that were tested. It also shows the percentage of the recalled prices that fell within 5, 10, and 20% of the actual prices. As one can see, the levels of price recall are quite different from those obtained in previous American studies. As a comparison, we know that in Dickson and Sawyer (1990), 56% of the respondents were within a 5% range of the actual price and 47% even recalled the exact price.

Table 1. Price Recall

| *Accuracy Level* | *Cumulative % of Respondents* |
|---|---|
| **Correct**[2] | 2.1% |
| **Within 5%** | 21.3% |
| **Within 10%** | 37.3% |
| **Within 20%** | 60.3% |

N = 1186 (missing values = 2.8%)

The difference between Dickson and Sawyer and our study can be due to the fact that we measure long-term memory while Dickson and Sawyer measure a mix of long and short-term memory. Another possibility is, of course, that French consumers have a much lower price knowledge than their American counterparts. To compare both explanations, we ran a separate survey, in the

same store, but now replicating the Dickson and Sawyer study as faithfully as possible on two product categories. Consumers were intercepted immediately after they placed a product in their cart. The categories selected were mineral water and yogurt. One hundred consumers were interviewed for each category. As shown in Table 2, the estimates of price recall accuracy are rather similar for the two categories and, more importantly, they are markedly different from those of Dickson and Sawyer. These data clearly indicate that the French consumers in our study paid less attention to prices and, as a result, were worse at short-term memory than their American counterparts.

To check for possible category biases, we considered the unit basis for which a price was given (respondents gave their estimates in the unit that came spontaneously to mind). Different unit responses were only observed for mineral water, milk, and toilet paper and 43% of the responses for these categories were expressed by item (bottle or roll), with the remainder stated by package. A *t*-test on absolute percent differences (abs[recall estimate - actual price]/actual price] shows that recall accuracy does not depend on the unit (per item = 22, per pack = 26; $t_{512} = 1.76$, p = 0.08).

Table 2. Replication of Dickson and Sawyer

| | *% correct** | *average error* | *% estimate* within 5%* | *% estimate* within 10%* | *% estimate* within 20%* |
|---|---|---|---|---|---|
| **Yogurt** | 10% | 15% | 30% | 54% | 70% |
| **Mineral water** | 10% | 20% | 27% | 48% | 66% |
| **Dickson & Sawyer** | 47% | 15% | 56% | | |

* percentage of respondents; n = 100 for each product category

**Price Recognition**

In our theory review, we stressed that it may be premature to draw conclusions for retail price policy by just looking at the accuracy of price recall. The other two dimensions of price

knowledge also have to be taken into consideration. Recognition of the actual price can be considered as the next step down on the price knowledge ladder. Table 3 gives the percentages of respondents that chose each response option. To correct for guessing and get an estimate of the level of genuine recognition, we apply a simple theory of guessing behavior that posits that a certain percentage of the respondents, *X*, recognize the actual price with perfect confidence, while the rest, 1- *X*, use pure guessing (Morrison 1981). Through random chance, one third of these guessers will pick the right answer. Thus, the number of observed correct responses (42.2%) is equal to $X + (1 - X) / 3$. Solving for $X$[3] we have a corrected percentage of non-guessers of 13.3%.

Comparing recognition with recall results provides some interesting insights. The percentage of correct recognition responses is obviously higher than that for recall (13.3% versus 2.1%). This recognition percentage can also be compared to the percentage of recall responses that were within 5%. Indeed, if someone were to give a price recall estimate that fell within 5% of the actual price, this person should, logically, be able to recognize the right price when it is presented alongside prices that are 10% higher and lower. Interestingly enough, only 63% of the respondents who recalled a price within 5% of the correct price were able to recognize the actual price. It looks like a sizeable number of consumers do not use their recalled price when they have to make a recognition judgment. A possible explanation for this result is that these people access different memory representations to answer the two questions and are not able to make the translation from one memory code for numbers to another. For example, they may have used a pictorial representation of the price to make their recognition judgment while they used the auditory verbal code to answer the price knowledge question. This is apparently a less effective strategy than if they had used the spontaneously recalled price as a basis for their recognition judgment as well.

Table 3. Price Recognition

| Level of presented price relative to actual price | % of Responses |
|---|---|
| **-10%** | 32.5% |
| **0** | 42.2% |
| **10%** | 25.3% |

N = 1186

**Deal Spotting**

For the price-alert shopper, recall and recognition are useful forms of knowledge but they suppose more or less precise representations of price in memory and reliable access to these representations. With our third form of questioning, we measured the extent to which our respondents only have a sense of magnitude of the normal price that guides them in judging which prices are attractive or unattractive. Our analysis combines the response patterns across the four deal spotting questions into one measure.

We start by dividing our responses into three groups. A first group is composed of consumers who are fairly knowledgeable about price magnitudes and not only respond positively to a large discount (and negatively to a large increase in price), but also respond in the right direction to small changes in prices (i.e., the answer "normal" or "good" for 5% decreases and "normal" or "bad" for 5% increases). This response behavior corresponds to a [good, good or normal, normal or bad, bad] response to a price series of [-20%, -5%, +5%, +20%], or a [bad, normal or bad, good or normal, good] response to a price series of [+20%, +5%, -5%, -20%]. Of all our responses, 32.7% fall into this category.

A second group is formed by consumers who do not even have the most minimal ability of deal spotting. It includes those who label the first presented price as good when it is actually 20% above the current price, and those who label a price reduction of 20% as a bad deal. We find

14.1% of our observations in this case. The remaining responses, that do not belong to the first or second group, in total 53.2%, reflect some intermediate form of deal spotting and constitute our third group. Overall, the really clueless segment of respondents is rather small (14.1%). Most of our sample (85.9%) is able to engage in some degree of deal spotting.

As in the case of the price recognition questions, some of the correct answers are the result of guessing. The observed number of responses that we considered as an indication of good deal spotting ability is therefore an overestimate of the true deal spotting ability. Correcting for guessing is unfortunately not as easy as for the price recognition question. We work in three steps (See Appendix B). In the first step we calculate the percentage of responses that would be categorized as correct deal spotting if in reality all respondents were using pure guessing. In the second step we infer, from our calculations of step one, an upper-bound estimate of the proportion of guessers in our sample. This then allows us to derive a lower-bound estimate of the number of true deal spotters. Our calculations give us an estimate of 26.9%. This is of course a worse case scenario. In the absence of perfect knowledge on the number of guessers, all we can say is that the true number of accurate deal spotters lies between 27 and 33%.

***Comparison of Forms of Price Knowledge***

In combination, our previous analyses of the three measures of price knowledge suggest the distinction between five steps on the price knowledge ladder (see Figure 2). At the first step there is no price knowledge. This is reflected in the absence of deal-spotting ability. The next step permits some level of magnitude sensitivity to *large* price differences but is not perfect. A third step helps one to *react* correctly to prices that depart from the usual price by at least 5%. Step four permits, in addition, accurate *recognition* of the actual price and the highest and rarest form of knowledge is the basis for accurate price *recall*.

The relative percentages of responses at steps two to five seem to confirm the intuition that deal spotting, price recognition, and price recall are incremental steps on the price knowledge ladder, meaning that some of those who have accurate deal spotting ability also have accurate recognition and that some of the latter group have accurate recall. To check whether the three memory measures are indeed steps of a uni-dimensional continuum, we ran a Guttman scalogram analysis (Robinson 1978). We calculated the coefficient of reproducibility of the data and compared it with the reproducibility under the assumption of independence among the measures. At 0.90, the observed reproducibility is not higher than the chance level of 0.92. A calculation of the coefficient of scalability (an indication of the extent to which the responses can be scaled on one dimension) confirmed this finding; at 0.23, it is clearly below the acceptable range of 0.60 to 0.65. In conclusion, although a comparison of the frequencies suggests that deal spotting, price recognition and price recall are memory measures of increasing difficulty along a single dimension, the Guttman scalogram analysis shows that instead there is more than one dimension at play. This may indicate that these measures tap into different dimensions of our memory system, a notion present in the literature on numerical cognition (Dehaene 1992). This finding is not only of importance to readers interested in memory phenomena. For those interested in reference-price effects, our results imply that consumers who exhibit reference price effects in choice models may have a form of price knowledge that allows them to react accurately to displayed prices but that does not permit recall in a survey.

***Drivers of Price Knowledge***

We test our hypotheses on price knowledge drivers by running logit regressions on our three measures of price knowledge (price recall, price recognition, and deal spotting) with the knowledge drivers as independent variables. To study price recalled, we use as binary dependent

variable whether responses are within a 5% range of accuracy or not[4]. For recognition, we look at whether the correct response was chosen or not (we cannot correct for guessing at an individual level) and for deal spotting, we consider the difference between the 32.7% of responses that demonstrate good deal spotting ability and the other responses. We also run an analysis in which we code, again with a binary variable, whether there was no deal spotting at all (14.1%) or some form of deal spotting (as with recognition, we cannot correct for guessing at an individual level). We expect answers of price oblivious shoppers to be completely random and therefore none of the independent variables should be significant.

To avoid multi-colinearity problems in our regressions, we run a factor analysis on the questionnaire items measuring the consumer-level knowledge drivers of hypotheses 4 to 6 (see Appendix C for details). This yields the three factors we expected: (1) in-store price information search, (2) across-store price information search, and (3) shopping trip budget. We use the factor scores for these three factors in our analysis rather than the actual survey answers.

In our questionnaire we also tried to assess the effect of memory recency ("When did you last purchase the category?") and salience ("Do you intend to purchase the category today?"). These questions came at the beginning of our questionnaire because we thought they could function as realistic memory cues for subsequent access to price knowledge.

In addition to our survey results, we have access to national panel data with the average brand loyalty indices per category to control for switching behavior, and the average purchase frequency for each category. A priori, the effect of switching behavior might go either way. On the one hand, if switching is motivated by price differences, frequent switching would indicate frequent processing of price information and therefore better knowledge. On the other hand,

frequent switching may also mean that consumers are exposed to many different prices, and thus it would be harder for them to remember any given price accurately. Regarding the other variables, one would expect consumers to remember prices more accurately for products they purchased frequently, recently, and for product categories that are more salient because they are on the current shopping list (a priming effect).

Shopper demographics such as gender and age were measured in the questionnaire and were used as control variables. However, they did not contribute to the model and were removed from the final analysis. The parameter estimates for the final model we estimated are reported in Table 4. The results give several important insights about consumers' knowledge of prices. First, none of the parameters in the *deal oblivion* regression are significant except for the intercept. This shows that consumers who are oblivious to prices are that way regardless of the characteristics of the product category and regardless of their individual characteristics and shopping behavior. Their knowledge is erratic and, as a result, their responses are random.

A second important finding is that, overall, the characteristics of product categories are significant predictors of price knowledge. As predicted, approximate price recall and recognition are worse for categories with a larger price range ($H_{2a,}$ $H_{2b}$) and more references ($H_{3a}$, $H_{3b}$). The results for price volatility on recall and recognition are also significant but go against our predictions: price promotions apparently not only make the promoted price but also the normal price better accessible to recall ($H_{1a}$) and recognition ($H_{1b}$). The hypotheses for the effect of price volatility ($H_{1c}$) and range ($H_{2c}$) are confirmed for deal spotting, but the parameter for category clutter is not significant ($H_{3c}$), though it has the expected sign (see Appendix D).

Table 4. Logit Regressions of Drivers of Price Knowledge

| | Accurate within 5% | | Price Recognition | | Deal Recognition | | Deal Oblivion | |
|---|---|---|---|---|---|---|---|---|
| | **Estimate** | **Odds Ratio** | **Estimate** | **Odds Ratio** | **Estimate** | **Odds Ratio** | **Estimate** | **Odds Ratio** |
| Intercept | -4.56** | - | -1.12** | - | -1.56* | - | -2.07* | - |
| Volatility **($H_1$)** | 1.17** | 3.21 | 0.49** | 1.63 | 0.58* | 1.78 | -0.36 | 0.69 |
| Range **($H_2$)** | -0.89** | 0.41 | -0.39* | 0.68 | -0.40* | 0.67 | 0.31 | 1.37 |
| Clutter **($H_3$)** | -1.32** | 0.27 | -0.62** | 0.54 | -0.23 | 0.80 | 0.31 | 1.37 |
| Frequency | 0.09** | 1.09 | 0.02* | 1.02 | 0.58* | 1.79 | -0.46 | 0.62 |
| Loyalty | 0.08** | 1.09 | 0.03** | 1.03 | 0.01 | 1.01 | 0.01 | 0.99 |
| Recency | 0.04 | 1.04 | 0.01 | 1.01 | -0.05 | 0.95 | -0.11 | 0.89 |
| Salience | 0.15 | 1.17 | -0.05 | 0.95 | -0.01 | 1.02 | 0.34 | 1.41 |
| In-store search **($H_4$)** (Factor 1) | 0.15* | 1.16 | 0.03 | 1.03 | 0.16* | 1.18 | -0.12 | 0.89 |
| Across-store search **($H_5$)** (Factor 2) | -0.09 | 0.91 | 0.09 | 1.09 | 0.02 | 1.02 | -0.01 | 1.00 |
| Shopping budget **($H_6$)** (Factor 3) | 0.04 | 1.04 | -0.02 | 0.98 | 0.05 | 1.05 | -0.01 | 1.00 |
| Chi-Square | p = 0.0001 | | p = 0.02 | | p = 0.0075 | | p = 0.32 | |

* Significant at p=0.05
** Significant at p =0.01

Although shopping frequency and brand loyalty are not measured at the level of the respondents but at the national level, these control variables are statistically significant. Frequency has the sign we had expected. Recall and recognition are better for categories that have higher than average levels of brand loyalty. Recency and salience did not turn out to be significant predictors of price knowledge. Thus, being in the store to buy the particular product does not by itself prime price knowledge and the price of more recently purchased products is not better accessible than that of products purchased in the more distant past.

The price search measures also give some important results. Probably the most surprising result is that people who claim to shop across different stores in order to "cherry pick" on the basis of price comparisons do not have better price knowledge, whatever the measure we consider. Also shoppers for large households or with big shopping budgets are not more knowledgeable about

prices, although they have most to gain financially from better knowledge. We therefore have to reject hypotheses 5a-5c and 6a-6c. Finally, hypothesis 4, that consumers who engage in in-store comparisons of prices have better knowledge is supported by our tests on recall ($H_{4a}$) and deal spotting ($H_{4c}$), but not for recognition ($H_{4b}$).

## Discussion of Results

### *Summary*

Our research clearly shows that a large majority of consumers hold some sort of price information for frequently purchased products in memory. Further, for most people these memorized prices are not an accurate representation of the last price seen, but rather a sense of magnitude. Either during the encoding of price information or in between shopping occasions, there is some information loss which makes it difficult for consumers to recall or recognize actual prices (<15% of success), but that is fortunately not so large that they are incapable of spotting good deals when they see them (>85% of success). This may indicate that, when confronted with hundreds of weekly purchase decisions, consumers develop a heuristic device for dealing with the vast amount of information to be processed, and that such a device provides them with the working knowledge necessary to make acceptable decisions (i.e., not pass good deals, nor be taken advantage of by large price increases). We also have some evidence that prices are represented in different forms in memory (sound sequences, photographic representations, senses of magnitudes) and that specific memory tasks require the ability to access the corresponding representation. This implies that, in order to get an accurate picture of price knowledge, future surveys should tap into the different forms of numerical memory through a combination of the appropriate questions.

Our analysis of the drivers of price knowledge indicates that frequent promotions make normal prices more memorable, which implies that when looking at promotions, consumers do pay attention to the *regular* price of the goods promoted, and can tell the difference between promoted and regular prices. A wider range between high-priced and low-priced items and a larger number of brands within a category seem to hamper price knowledge, possibly because these factors increase the complexity of the information that is to be remembered. This may imply that in terms of memory organization, prices are not just linked to the respective brands but somehow are also related to the product category. From a learning perspective, it is no surprise that more frequently purchased product categories and more loyalty to a brand lead to better price knowledge.

All the previous findings are related to the product category, which means that they can be directly applied by retailers for their pricing strategies. This is not the case for shopper characteristics where retailers first have to identify which type of customer they have in the store in order to be able to adapt their price policy. For these shopper characteristics, our most surprising result may be that the practice of cherry-picking has no impact on price knowledge. This can be tentatively explained by the increase in task complexity for cherry-pickers. For each product in their shopping basket they have to keep either a different reference price for each of the stores they patronize, or they have to create a reference price from the aggregate of multiple sources of prices, which would likely reduce the precision and the actionability of their price knowledge. A final finding on price knowledge drivers is that recency does not affect price knowledge, which supports the claim of Briesch et al. (1997) that reference prices are generated over extended periods of time.

***Implications for Retailers***

Previous research has warned managers in the retailing industry against their tendency to overestimate the percentage of consumers who search and respond to price information (e.g., Urbany, Dickson, and Kalapural 1996). We find that cherry-pickers for frequently purchased goods do not pose a major threat because their price knowledge is not different from that of other shoppers. Overall, shoppers demonstrate low accuracy in price knowledge, but they have the ability to detect attractive and unattractive prices. This lack of precise knowledge helps explain the results obtained by Hoch, Drèze, and Purk (1995) in their study of Every Day Low Pricing (EDLP) vs. Hi-Lo pricing. In their paper, they showed that a change of plus or minus 10% on regular prices had little impact on store sales. This is consistent with our results since our analysis indicates that only a small minority of consumers has the information necessary to notice such a small change in regular non-promoted prices. Hence, consumers would be hard pressed to notice the change in regular price, but we may assume they would still be able to notice the promotional activity. In short, this suggests that consumers don't know what *Hi* is, but they can recognize *Lo* when they see it. In other words, as long as *Hi* does not increase to the point at which it can be recognized as being a bad price, sales will not be affected.

In addition to explaining the lack of performance of the EDLP format in Hoch *et al* (1995), our findings are relevant to the work on inter-store competition. The important points here are that we did not find consumers to be proficient at remembering regular prices, but we found them to be proficient at spotting deals. Further, they are better at remembering regular prices when prices are promoted often than when prices are promoted infrequently. This indicates that it will be difficult for two EDLP stores to compete on price since few consumers have the information necessary to make a valid store comparison. On the other hand, two Hi-Lo stores can compete through

promotions. Further, in a store format comparison, the credibility of EDLP prices is heightened by the presence of a Hi-Lo store as the Hi-Lo store makes regular prices more salient.

In terms of in-store pricing, our findings indicate that consumers' knowledge of price is category-dependent. The accuracy of knowledge depends on such factors as category clutter (i.e., the number of brands in a category), volatility (the frequency of promotions), and price range (the price difference between the highest priced item and the lowest priced one). It follows that a retailer's pricing strategy should also be category dependent. To get the image of offering low prices, a retailer should first focus on categories that facilitate price knowledge (low clutter, small price range, and frequent promotions) should have lower prices relative to categories that make it more difficult to memorize and compare prices (e.g., categories with high level of clutter, large price ranges, and infrequent promotions). The former group of categories will attract price-sensitive consumers to the store, while the latter will permit the retailer to maintain a reasonable overall margin.

***Price Knowledge and Reference Price***

In the introduction we referred to reference prices in the context of choice models. This stream of research conceptualizes reference price as the price knowledge residing in long-term memory resulting from exposure to past prices. Reference price is, in the absence of direct measures of price memory, operationalized in choice models as some average of the history of actual prices. Looked upon in this context, the measures of price awareness used in the present study can be viewed as an operationalization of reference price and are a step forward in the direct measurement of reference price. However, what we lack in our survey is the act of choice and we can therefore not claim that our measures of price awareness are also measures of reference price. We do not know if consumers would use the price they give in a recall or recognition question to

make an actual price comparison. In contrast, the deal-spotting questions clearly require the interviewee to make a comparison between an observed price and some internal benchmark. The deal spotting question can therefore be considered as a direct measure of reference price, be it that it remains a partial measure because it taps mainly into one of the three forms of memory for numbers, namely the analogue magnitude code.

Although the link between our measures of price knowledge and the concept of reference price remains to be validated, our research makes several contributions to the debate between reference price and price survey studies. First, it is obvious now that price surveys based on short-term recall reflect attention to prices and therefore only have very *indirect* implications for reference price. They will in most cases overestimate the level of long-term price knowledge in the market place. Second, price surveys on long-term recall severely underestimate the amount of actionable consumer price knowledge, because they do not take memory representations that are only accessible to recognition and recall into account. Choice models can therefore detect reference price effects even if price recall is inaccurate or unreliable. Although further validation is required, the two previous points can be considered as a reconciliation between previous price surveys, that show low price knowledge, and reference price models that suggest a much higher price knowledge. Third, reference price models that focus on reference price zones instead of reference price points (e.g., Kalyanaram and Little, 1994) represent more valid descriptions of the way the majority of consumers hold price information in memory, because price zones represent deal spotting ability, a form of price knowledge that is widely available in the market, and correspond to the magnitude code representation.

***Limitations***

Before concluding this paper, we must be cognizant of the shortcomings of the method we used. Our basic instrument is a survey that was administered at the start of a shopping trip and therefore had to be as time-efficient as possible. Some factors that probably are important determinants of price knowledge were not included in the survey to limit its length (e.g. individual-level rather than market-level brand loyalty). In addition, environmental variables such as the level of market share, promotion, and advertising of the brands used in our study were not included in our analysis. We clearly only examined a small subset of the relevant knowledge drivers.

A second limitation concerns the fact that each respondent answered, for a given product, to a recall question, four deal-spotting questions, and a recognition question. This setup was essential to examine to what extent the three corresponding forms of numerical memory are dependent or independent. It has, however, the drawback that carry-over effects from one question to another are possible, although the sequence of questions that we adopted was the one that minimized these carry-over effects. The most important concern is that our estimate of recognition may be inflated. Respondents saw a sequence of four prices for deal spotting and the average or midpoint of these prices (which was not stated itself) was the correct answer for the recognition question. Moreover, this correct answer was also the midpoint of the prices presented in the recognition question. It is therefore possible that some respondents figured out our design, even though they did not get any feedback on the accuracy of their responses.

This is an important concern and we examined it ex-post in a new survey. In order to get an upper-bound estimate of the possible bias of our design, this new survey was designed to maximize our respondents' chance of figuring out the logic in our price structure. We administered our price knowledge questions to a group of people who could not have any idea of

the actual prices of the products and therefore had to rely entirely on logic to determine the best possible answer. One hundred and five American college students had to answer the deal-spotting and recognition questions for a 4-pack of Danone Fruit Yoghurt, a 6-pack of 1.5l bottles of Evian, and 1kg of sugar. The prices were stated in French francs (the students were not told what the current exchange rate between the US dollar and the French franc was) and the question pertained to French supermarkets ("is the price shown a good, normal, or bad price for 1kg of sugar in a typical French supermarket?"). The actual survey instructions and the analysis of the answers are shown in Appendix E. It is clear that the answers from the store survey are significantly better than the results of the pure guess survey ($p=0.0001$). This indicates that the results obtained in the main survey are not due to logic-based inferences alone. It does not, however, rule out the fact that some guessing occurs, which is why we apply our correction factor to the store survey results for recognition, scaling down the correct response rate from 42.2% to 13.3%.

The fact that we restricted our sample to consumers who made most of their shopping trips at the participating retailer is a third limitation of our study. We did so to increase the validity of our comparison of the remembered prices to the actual store prices but it would be interesting to record prices at a variety of stores and conduct the same study with a group of known cherry-pickers to test whether they keep separate prices in memory for the separate stores. We will leave an in-depth analysis to future research, but nevertheless attempted to verify the impact of this factor on our results. One of our most surprising findings, that respondents who report across-store price comparisons do not have a better price knowledge than those who do not, could be due to our sample selection in that we may have insufficient variance in across-store shopping. This is not the case: with 19% of the respondent reporting that they regularly shop in two stores, 19% in three and 22% in more than three stores, lack of variance cannot be a concern. A second

alternative explanation may be that the effect of cherry-picking as such was attenuated by the effect of the other items that constituted the across-store search factor. We therefore reran all our logit regressions and replaced factor 2 first by the reported number of stores visited and then by the item that measured cherry picking. Neither of these variables had a significant effect. In conclusion, the null-effect of across-price search cannot be explained by biases introduced by our survey procedure or our analysis.

A final limitation to keep in mind is that we estimated price knowledge for the UPC a respondent buys most often in a product category. Our results should therefore not be generalized to the respondent's entire shopping basket.

## Conclusion

We showed in this study that the accuracy of consumers' price knowledge depends on both the shopping environment (e.g., category clutter, promotion activity) and the idiosyncrasies of consumers (e.g., brand loyalty, in-store price search behavior). These findings are important to managers who have to make pricing decisions in a competitive multi-product environment. In addition, our distinction between the three forms of number representation in memory and our combination of the three corresponding measures (recall, recognition, deal spotting) explain the apparent contradiction between the observations made in price knowledge surveys (consumers have low levels of price knowledge) and those in reference price studies on scanner data (consumers' decisions indicate a high level of price knowledge). Although most consumers do not possess an accurate knowledge of price that permits accurate recall or recognition, they possess a working knowledge of prices that is accurate enough to make good purchase decisions.

This working knowledge, and not accurate price recall as such, may then be the driving force behind the reference price effects observed in choice models.

## References


Anderson, Michael C. and Bjork, Robert A. (1994), "Mechanisms of Inhibition in Long-Term Memory: A New Taxonomy," *Inhibitory Processes in Attention, Memory, and Language,* in D. Dagenbach and T. Carrs (Eds*.)* (p. 265-325). San Diego: Academic Press.

Ashcraft, Mark H. (1992), "Cognitive Arithmetic. A Review of Data and Theory," *Cognition*, 44, 75-106.

Briesch, Richard A., Lakshman Krishnamurthi, Tridib Mazumdar, and S.P. Raj (1997), "A Comparative Analysis of Reference Price Models," *Journal of Consumer Research,* 24 (September), 202-214.

Dehaene, Stanislas (1992), "Varieties of Numerical Abilities." *Cognition* 44:1-42.

---- and Rokny Akhavein (1995), "Attention, Automaticity, and Levels of Representation in Number Processing," *Journal of Experimental Psychology,* 21 (2), 314-326.

Dickson, Peter R. and Alan G. Sawyer (1990), "The Price Knowledge and Search of Supermarket Shoppers," *Journal of Marketing*, 54, July, 42-53.

Hoch, Stephen J., Xavier Drèze and Mary E. Purk (1994), "EDLP, Hi Lo and Margin Arithmetic," *Journal of Marketing* 58, 16-27.

Johnson, H.M. (1994), "Processes of Successful Intentional Forgetting," *Psychological Bulletin*, 116, 274-292.

Kalyanaram, Gurumurthy and John D.C. Little (1994), "An Empirical Analysis of Latitude of Price Acceptance in Consumer Package Goods," *Journal of Consumer Research*, 94 (December), 408-418.

---- and Russel S. Winer (1995), "Empirical Generalizations from Reference Price Research," *Marketing Science*, 14 (No. 3, Pt. 2), G161-G169.

Le Boutillier, John, Susanne Shore Le Boutillier and Scott A. Neslin (1994), "A Replication and Extension of the Dickson and Sawyer Price-Awareness Study," *Marketing Letters*, 5, 1, 31-42.

Mandler, George (1980), "Recognizing: The Judgment of Previous Occurrence," *Psychological Review,* 87, 3, 252-271.

McCloskey, Michael and Paul Macaruso (1995), "Representing and Using Numerical Information," *American Psychologist*, 50 (May), 351-363.

Monroe, Kent B. and Angela Y. Lee (1999), "Remembering Versus Knowing: Issues in Buyers' Processing of Price Information," *Journal of the Academy of Marketing Science*, 27, 2, 207-255.

---- , Christine P. Powell, and Pravat K. Choudhury (1986), "Recall versus Recognition as a Measure of Price Awareness," in *Advances in Consumer Research*, Vol. 13, ed. Richard Lutz, Provo, UT: Association for Consumer Research, 594-599.

Morrison, Donald G. (1981), "Triangle Taste Tests: Are the Subjects Who Respond Correctly Lucky or Good?," *Journal of Marketing*, 45 (Summer), 111-119.

Que Choisir (1998), "Près de Chez Vous, la Grande Surface la Moins Chère," Juin, 30-41.

Robinson, John P. (1978), "Toward a More Appropriate Use of Guttman Scaling," *Public Opinion Quarterly*, 42, 2, 235-246.

Urbany, Joel E. and Peter R. Dickson (1991), "Consumer Normal Price Estimation: Market Versus Personal Standards." *Journal of Consumer Research,* 18 (June), 45-51.

----, ---- and Rosemary Kalapurakal (1996), "Price Search in the Retail Grocery Market," *Journal of Marketing*, 60 (April), 91-104.

Wakefield, Kirk L. and J. Jeffrey Inman (1993), " Who are the Price Vigilantes? An Investigation of Differentiating Characteristics Influencing Price Information Processing." *Journal of Retailing,* 69 (Summer), 216-233.

Winer, Russel S. (1986), "A reference price model of brand choice for frequently purchased products," *Journal of Consumer Research*, 13 (September), 250-6.

## Appendix A: Formal Hypotheses

$H_{1a}$ The ability of a consumer to *recall* the price of an item is negatively affected by the level of price volatility in the category in which the product belongs.

$H_{1b}$ The ability of a consumer to *accurately recognize* the price of an item is negatively affected by the level of price volatility in the category in which the product belongs.

$H_{1c}$ The ability of a consumer to *spot a good or bad deal* for a product (i.e., recognize that a price cut of 20% is a good deal, or that an increase of 20% is a bad deal) is positively affected by the level of price volatility in the category in which the product belongs.

$H_2$ The ability of a consumer to *recall* a price ($H_{2a}$), *accurately recognize* a price ($H_{2b}$) and *spot a good or bad deal* on a product ($H_{2c}$) are negatively affected by the range of prices in the category to which the product belongs.

$H_3$ The ability of a consumer to *recall* a price ($H_{3a}$), *accurately recognize* a price ($H_{3b}$) and *spot a good or bad deal* on a product ($H_{3c}$) are negatively affected by the level of clutter in the category in which the product belongs.

$H_4$ Consumers engaging in in-store price research will be better at *recalling* ($H_{4a}$) and at *accurately recognizing* ($H_{4b}$) the price of an item and at *spotting a good or bad deal* on a product ($H_{4c}$) than those who don't.

$H_5$ Consumers engaging in across-store price search will be better at *recalling* ($H_{5a}$) and at *accurately recognizing* ($H_{5b}$) the price of an item and at *spotting a good or bad deal* on a product ($H_{5c}$) than those who don't.

H6 Consumers who have large average basket sizes will be better at recalling ($H_{6a}$) and at accurately recognizing ($H_{6b}$) the price of an item and at spotting a good or bad deal on a product ($H_{6c}$) than those who don't.

# Appendix B: Deal Spotting Guessing Corrections

### *B.1 Probability of Guessing Right*

To correct for guessing in the deal-spotting question, we begin by assuming that a guesser will answer at random, but will answer in a logical manner. That is, a guesser will not decrease his/her valuation of a deal (e.g., go from 'Normal' to 'Bad') when presented with a price that is lower than the previous price. Similarly the guesser will not better his/her valuation of a deal (e.g., go from 'Normal' to 'Good') when presented with a price that is higher than the previous price.

Hence, if presented a series of 4 increasing prices [-20%, -5%, +5%, +20%], a guesser who started with 'Bad' would be compelled to rate all four prices as 'Bad.' In contrast, a guesser who started with 'Good,' could answer any of the ten following sequences:

| | | | | |
|---|---|---|---|---|
| 1. | Good | Good | Good | Good |
| 2. | Good | Good | Good | Normal |
| 3. | Good | Good | Good | Bad |
| 4. | Good | Good | Normal | Normal |
| 5. | Good | Good | Normal | Bad |
| 6. | Good | Good | Bad | Bad |
| 7. | Good | Normal | Normal | Normal |
| 8. | Good | Normal | Normal | Bad |
| 9. | Good | Normal | Bad | Bad |
| 10. | Good | Bad | Bad | Bad |

Any other answer (e.g., Good, Normal, Good, Bad) would be illogical and thus will not be considered in the analysis. Of these 10 possible guessing sequences, only 4 (sequences 5, 6, 8, and 9) would be considered to be correct answers for the deal spotting questions. Guessing sequences 1, 2, 4, and 7 are incorrect because they fail to recognize the +20% price as 'Bad;' guess 3 is incorrect because it classifies +5% as a 'Good' deal; guess 10 is incorrect because it classifies –5% as a 'Bad' deal.

To compute the probability of guessing sequences 5, 6, 8, or 9, we need to recognize that the probability of guessing 'Good,' 'Normal,' or 'Bad' depends on the choice made on the previous

answer. ‘Good’ will be guessed with a 1/3 probability on the first guess or if the previous guess was ‘Good;’ it will be guessed with probability 0 if the previous guess was ‘Normal’ or ‘Bad’ (in an increasing sequence). Similarly, ‘Normal’ will be guessed with probability 1/3 on the first try or if the previous guess was ‘Good,’ with probability ½ if the last guess was ‘Normal,’ and 0 if the previous guess was ‘Bad.’ Finally, ‘Bad’ will be guessed with probability 1/3 on the first try or if the previous guess was ‘Good,’ ½ if the previous guess was ‘Normal,’ and 1 if it was ‘Bad.’

We can now compute the probability of sequence 5 [Good, Good, Normal, Bad] as 1/3 * 1/3 * 1/3 * 1/2 = 1/54. Similarly sequence 6 is guessed with probability 1/3 * 1/3 * 1/3 * 1 = 1/27 and the result for 8 is 1/36 and that for 9 is 1/18. Hence, the probability of a correct guess is 1/54 + 1/27 + 1/36 + 1/18 = 13.9% as illustrated graphically on figure B1.

***B.2 Probability of being deal oblivious***

To be classified as deal oblivious, one must answer ‘Bad’ for the first price in an ascending sequence, or ‘Good’ for the first price in a descending sequence. As explained in the preceding section, the probability of either of these answers is 1/3. Given that half the series were ascending and the other half were descending, the probability for a guesser to be classified as deal oblivious is 1/3 * 1/2 + 1/3 * 1/2 = 1/3.

***B.3 Probability of reflecting intermediate deal spotting abilities***

We classify as intermediate all shoppers who are neither correct nor deal oblivious. As was shown in the preceding two sections, the probability of being correct is 13.9% and the probability of being oblivious is 33.3%. This leaves 52.8% probability of being intermediate.

### *B.4 Upper and lower bounds on deal spotting abilities*

If *everybody* were guessing, we would see 13.9% of correct answers, 33.3% of deal oblivion, and 52.8% of intermediate answers. We observed 32.7%, 14.1%, and 53.2%. These numbers are different enough to rule out the possibility that everybody was guessing. But, how many people truly knew? The observed 32.7% is obviously an upper bound on the percentage of people who can truly spot a deal accurately as some of our respondents must have hit on the right answers by chances. Indeed, 13.9% of the guessers would have done so. The question is, what is the lower bound? To figure this out, we need to determine an upper bound on the number of guessers. The upper bound on guessers is 42%. Indeed, if 42% of the people guessed, they would yield 42% * 13.9% = 5.8% of correct answers, 42% * 33.3% = 14.1% of deal oblivion, and 42% * 52.8% = 22.2% of intermediate answers. If there were more than 42% of guessers, we would see more than 14.1% of deal oblivion, which we did not see. Hence, 42% is the upper bound on guessing. This would yield 5.8% of correctly guessed answers. Hence, 32.7% - 5.8% = 26.9% is our lower bound on the true number of accurate deal spotters.

## Appendix C: Factor Analysis of Consumer and Shopping Characteristics

Rotated factor pattern after Varimax rotation

| | *Factor 1 In-Store Price Search* | *Factor 2 Across-Store Price Search* | *Factor 3 Shopping Trip Budget* |
|---|---|---|---|
| **Do you pay attention to in-store promotions?** | **0.75** | 0.12 | 0.13 |
| **Do you compare the flyers you find at the entrance of the store or in your mailbox?** | **0.71** | 0.29 | 0.07 |
| **Do you like shopping at supermarkets?** | **0.57** | -0.18 | -0.28 |
| **How often do you shop at different stores to buy at the best possible price?** | 0.30 | **0.75** | -0.05 |
| **Do you compare prices between different stores?** | 0.47 | **0.64** | -0.01 |
| **In how many supermarkets do you do your weekly shopping?** | -0.21 | **0.79** | 0.03 |
| **How many members are there in your household?** | 0.10 | 0.02 | **0.80** |
| **How much do you spend on average in this store?** | -0.07 | 0.09 | **0.80** |

Notes: the items appeared in a different order in the questionnaire.

Loadings in boldface indicate the questions that load the highest on each factor.

Appendix D: Summary of Hypothesis-Testing Results

| | ***Effect on Price Recall*** | ***Effect on Price Recognition*** | ***Effect on Deal Spotting*** |
|---|---|---|---|
| ***H1: Category Volatility*** | *Reverse* | *Reverse* | *Support* |
| ***H2: Price Range*** | *Support* | *Support* | *Support* |
| ***H3: Category Clutter*** | *Support* | *Support* | No Support |
| ***H4: In-store price research*** | *Support* | No Support | *Support* |
| ***H5: Across-store price research*** | No Support | No Support | No Support |
| ***H6: Average Shopping Trip Size*** | No Support | No Support | No Support |

## Appendix E: Test of Carry-Over Effects: Survey among Colleges Students

When asking the deal spotting questions we used series of prices that were either [-20%, -5%, +5%, +20%] or [+20%, +5%, -5%, -20%] off the regular prices. When asking the price recognition question, we used the regular prices along with prices that were 10% below and 10% above regular prices. The combination of these two questions may allow an astute interviewee who does not know prices to guess that the regular price is the middle price in the second question.

To test for this possible bias in the results due to guessing, we ran the same questions using American undergraduate students from a major West coast school. That is, we ask American students questions about French prices in French Francs. Students were not told what the exchange rate from the Dollar to the Franc was. Further, this questionnaire was run on the last week of the fall semester (15 weeks) to ensure that students would have to guess the answer (it is unlikely that these students would have had a recent shopping experience in France).
The survey started by asking the deal-spotting question for a 4PK of Danone Fruit Yogurt, starting with a price that was 20% below regular price and increasing to 20% above regular price. Students were then asked the price-recognition question for the same product. The questions were repeated for a 6PK of 1.5 Liter Bottles of Evian but this time with a decreasing price sequence for the deal-spotting question. Finally, we concluded with the deal-spotting (decreasing sequence) and price recognition questions for 1Kg of Sugar.

The price recognition questions all used prices that were 10% above, 10% below, and at regular price. However, the order of the prices where rotated using a Latin-square design as follows:

Table E1. Latin-square design of the ex-post survey with product category (in columns), presentation position (in rows) and price level (in cells)

| | Danone | Evian | Sugar |
|---|---|---|---|
| Left | Regular | -10% | +10% |
| Middle | +10% | Regular | -10% |
| Right | -10% | +10% | Regular |

Interestingly enough, the price that was picked the most often was not the middle price in terms of franc value, but rather the price that was in the middle visually (as opposed to the price on the left or the right). This center price was chosen in 41% of the cases (significantly different from random choice at p=0.002). This 'center' bias does not concern us for the store survey as we took care to rotate the order of the prices in the store survey.

A comparison of the results from both surveys shows that the answers from the main survey are significantly better than the results of the pure guess survey (p=0.0001). This indicates that the results obtained in the main survey are not due to guessing alone. It does not however rule out the fact that some guessing occurs, which is why we apply our correction factor to the survey results, scaling down the correct response rate from 42.2% to 13.3%.

Figure 1: Category Selection

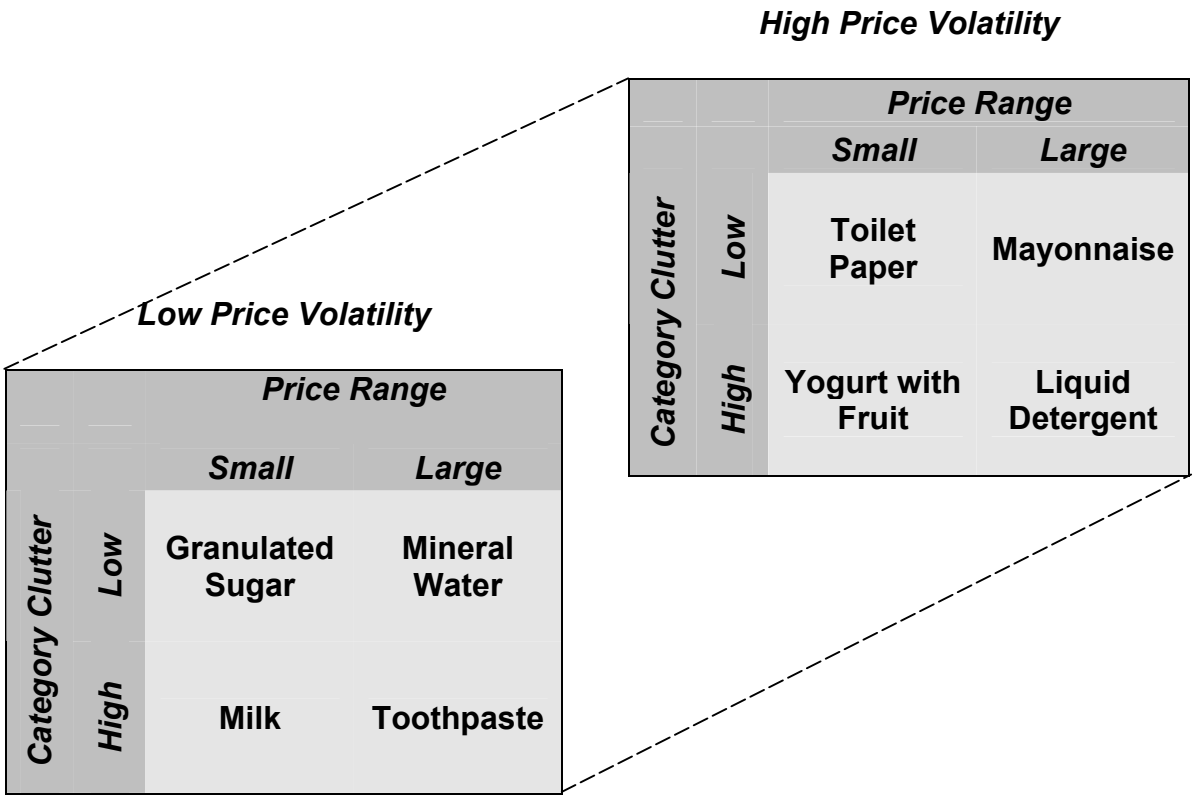
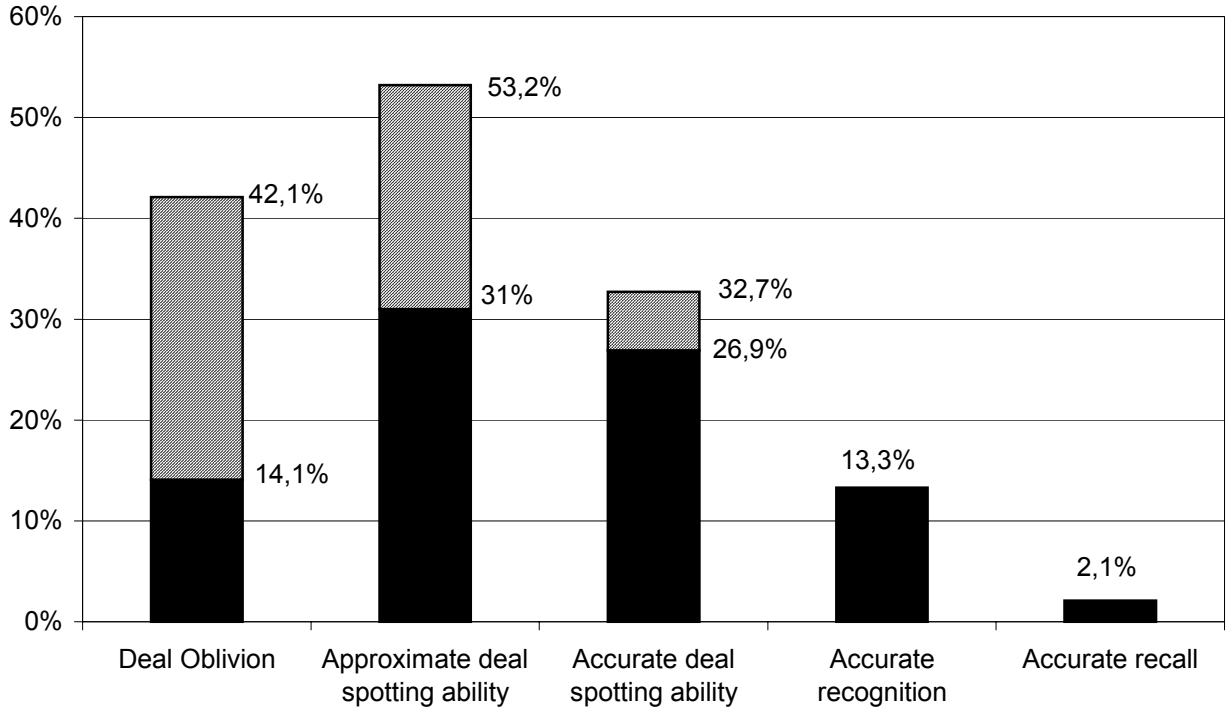
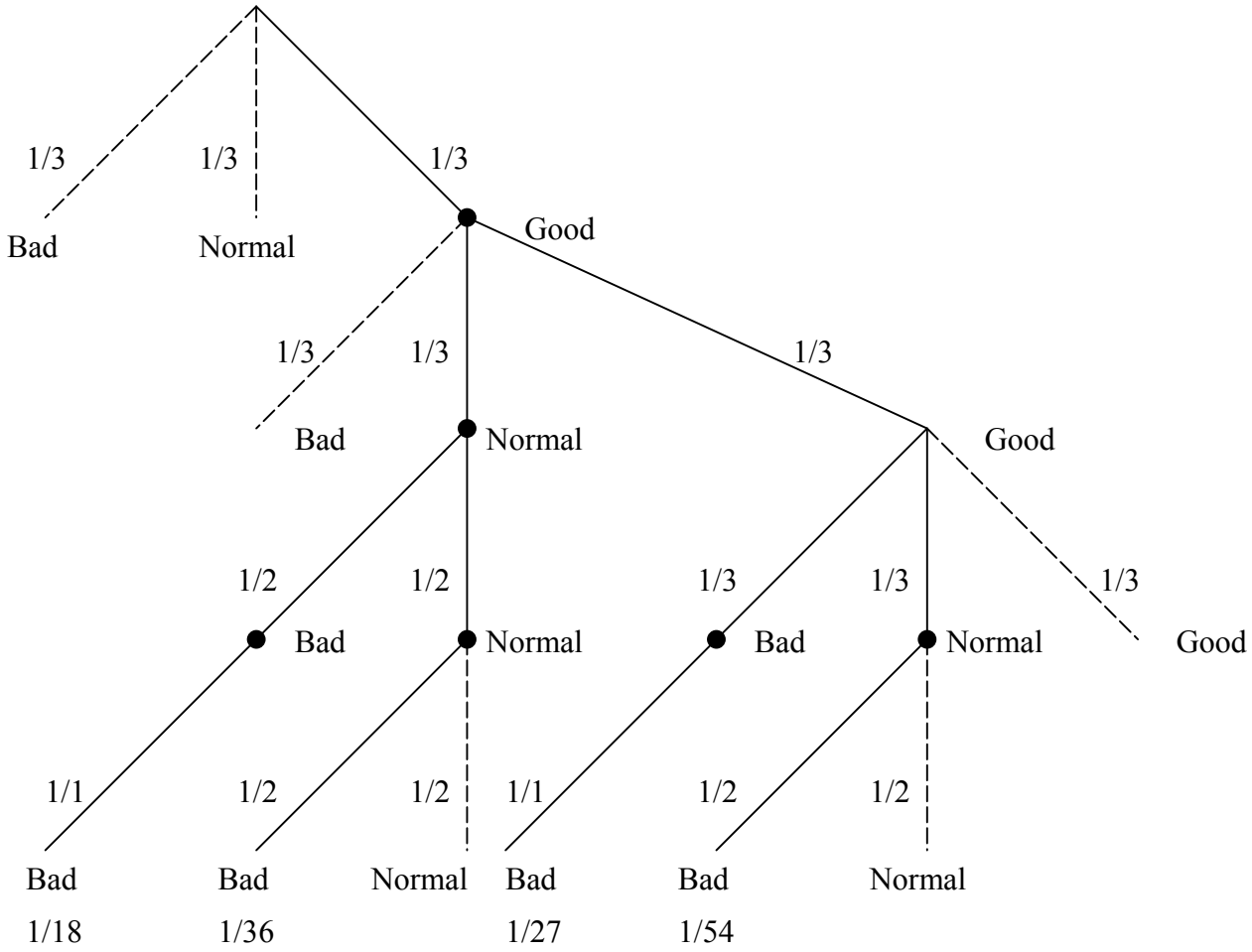


Figure 2: Levels of Price Knowledge




**Legend for Figure 2.** The bar heights and percentages represent independent estimates and therefore do not add up to 100%. For deal spotting both the upper and lower bounds of our estimates are given. The upper bounds are equal to the observed proportions of the three levels of deal spotting ability. The lower bounds correspond to the assumption that the number of guessers in the sample reaches its maximum.

Figure 4: Guessing probabilities for an increasing sequence




Probability of correct guess: 1/18 + 1/36 + 1/27 + 1/54 = 13.9%

## Footnotes

[1] This financial incentive may have introduced a selection bias by attracting price sensitive shoppers. This bias does however not affect the nature of our conclusions, as will become obvious in the analyses.

[2] We defined accuracy relative to the current price in the store. This may not always correspond to the reference price, for instance if the reference price is based on several past prices and when the price changed from the previous shopping occasion. It should, however, also be noted that prices in this store are fairly stable and current price is therefore a good approximation of the "correct" reference price.

[3] $0.422 = X + (1 - X) / 3 \Rightarrow X = (3 * 0.422 - 1) / 2 = 0.133$.

[4] Although we only report one analysis of the recall results, we compared different approaches. The results are not substantially different. For instance, a first analysis took absolute percentage differences as dependent variable. When all data were included, we obtained identical results as those reported with one difference: range was no longer significant. This result was, however, influenced by a number of extreme outliers. When excluding the 1% most extreme outliers there are no differences with the reported results.