Do Consumers Really Know if the Price Is Right? Direct Measures of Reference Price and Their Implications For Retailing

Marc Vanhuele
Xavier Drèze*

September 2000

*Marc Vanhuele is Associate Professor of Marketing at HEC Paris (France), and Xavier Drèze is Assistant Professor of Marketing at the University of Southern California, Los Angeles. The authors thank the HEC Foundation for its financial support, and Shantanu Sutta and Gilles Laurent for their comments on a previous draft.

Abstract

Reference price research suggests that consumers memorize and recall price information when selecting brands for frequently purchased products. In this study, we show that previous price-knowledge surveys provided imperfect estimates of reference price. Further, we propose 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. In addition we identify consumer and product characteristics that explain the variations in price knowledge.

There exists a large body of empirical evidence showing that, when making brand choices for packaged consumer goods, consumers compare observed prices to so-called *reference prices* they supposedly have in memory (cf., Winer 1986, or Kalyanaram and Winer 1995 for a review). Different models of reference price have been validated in the past and, in a comparison of the most common model formulations, Briesch et al. (1997) found that the best performing model is one that includes brand-specific reference prices, represented as a moving average of the price history of each brand in a category. This suggests that consumers store brand-specific price information in long-term memory, and update this information after each purchase occasion. The evidence for the existence of memory-based reference prices is nevertheless *indirect.* Choice models based on panel data indicate that in their brand selection, consumers act *as if* they have access to and use this type of information, but in reality, reference price knowledge is not measured directly. Reference price models imply that (1) consumers notice current prices, (2) they access the reference price in long-term memory, and (3) they process both prices at a sufficient level to allow comparisons. These assumptions are put into question by direct surveys of price knowledge.

Whereas reference price studies based on panel data infer the use of reference prices without ever measuring their existence, another stream of research has focused specifically on measuring the level of price knowledge among consumers. Studies that directly measure price knowledge (Dickson and Sawyer 1990; Le Boutillier, Le Boutillier, and Neslin 1994; Wakefield and Inman 1993) show that a surprisingly low percentage of consumers pay attention to prices, even while shopping. The numbers reported are low especially compared to what reference price research suggests (see Monroe and Lee 1999 for a comparison). Only 47 to 55% of the respondents can accurately recall the price of an item they have just placed in their shopping cart. 19 to 23% do

not even attempt to give an estimate. In reaction to these results, several authors have 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). The same results have, however, also been interpreted as an indication that "a sizeable proportion of consumers are capable of recalling prices" (Briesch et al. 1997, p. 213; also see Kalyanaram and Winer 1995). In this paper we argue that, in fact, neither interpretation is satisfactory because the price measured in the above surveys does not correspond to the concept of reference price.

Indeed, studies such as Dickson and Sawyer or their replications do not tap into consumers' long-term memory. Rather, they ask price recall questions immediately after the consumer picked a product from the shelf. Although, in theory, a respondent can produce a response by consulting long-term memory, s/he is more likely to retrieve an answer from short-term memory. Price surveys therefore probably mainly measure the fact that the current price was noticed and retained in *short-term* memory. Without rehearsal or elaboration on this price information, it is unlikely to be transferred to long-term memory. Reference price is, in contrast, defined as a form of *long-term* memory because it is supposed to represent prices from previous shopping occasions. Because of the difference between long-term and short-term memory, the direct estimates of price recall currently available in the literature tell us very little about the existence and level of reference price. Using short-term price recall as a measure of reference price would considerably *overestimate* the extent to which actual reference prices exist. This further accentuates concerns raised in the past on the apparent contradictory evidence between choice models and direct questioning of price knowledge. Interestingly, however, one could also argue that, in theory, measuring price memory with a recall question will lead to an *underestimation* of

the actual level of knowledge. Instead of relying on recall, consumers may be alerted to price changes by recognition (or more precisely, the *lack of recognition*). Still another possibility is that there exist forms of sensitivity to price magnitudes that are not accessible to either recall or recognition. Reference price effects may then be caused by forms of memory that are not reflected in traditional recall and recognition measures.

The starting point of this paper was the apparent contradiction between two streams of research on price knowledge: one on indirect measures in the form of reference price through choice modeling (e.g., Winer 1986) and one on direct memory measures (e.g., Dickson and Sawyer 1990). To examine this contradiction, the first objective of our study is to propose a set of direct price knowledge measures that are designed to capture any form of knowledge that can account for the reference price effect. Recallable knowledge, as measured by Dickson and Sawyer, is only one such form. Our second objective is to study the consumer and environmental factors that can influence the ability of shoppers to store and use price knowledge. By looking at the environmental and behavioral factors affecting price knowledge, we are able to characterize differences in price knowledge across consumers and across categories that are important to managers who have to coordinate prices in a competitive multi-category environment. Because of the combination of these two objectives, our study is of importance to marketers and academics interested in the role of price knowledge and price comparisons for low-involvement purchases.

Some of our findings may appear counter-intuitive at first (for instance consumers who compare prices across stores do not have a better knowledge of price than consumers who do not), they are

nevertheless important to managers who have to coordinate prices in a competitive multi-category environment.

In the next section we examine, from a theoretical perspective, how consumers store prices in their memories and, consequently, which measures should be used to determine the level of price knowledge. This is then used to develop a survey instrument to measure the different dimensions of long-term consumer price knowledge that may account for reference price effects. The survey was run with 400 respondents in a French supermarket. The analysis of the results first treats the different knowledge dimensions separately then examines to what extent they are related. Next, the factors that drive price knowledge are analyzed. The final section of the paper then discusses the implications for price setting and for future research in this area.

Long Term Memory and Number Representation

In order to understand how consumers develop and process reference prices, we must understand how numbers are stored in long-term memory. Numerical cognition literature identifies three different types of memory representations of price. This research 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 proposing that numbers can be mentally represented and manipulated in three different forms, depending on the task at hand. 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). Each representation has its input-output procedures, although conversions from one code to another are possible. Important for our study is that Dehaene posits that each number-related task can be decomposed into a sequence of processes that each require a specific form of input. Different individuals can also have different preferred entry codes. Any of the three codes can be involved in a reference price comparison. For example, a verbal recall question requires accessing of the auditory verbal code. For a recognition question, access to the visual Arabic code is sufficient, while estimates of the normal price range can be based on the magnitude code. By using each of these types of questioning we can take the three possible forms of representation of price knowledge into account.

The three forms of measuring price knowledge have also been discussed individually in marketing literature. Although recall has been used most often in past studies, Monroe et al. (1986) have argued that recognition is a more appropriate measure than recall for price knowledge. 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 more 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 be more appropriate as a price knowledge measure than recall.

Regarding the notion that looking for price magnitudes instead of precise price points may in some cases be more appropriate, there is research that suggests 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.

Towards a Direct Measure of Reference Price

Based on our literature review, we propose that an estimation of reference price knowledge needs to tap into all three forms of price knowledge. Individually or in combination, these three forms of long-term price knowledge can produce the reference price effect of a choice model. In our survey of consumers we then combine the three measures that correspond to the 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. It can consist of a price point estimate or of a price range (if the verbal code is only partial).
- *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. Even without further information they can then infer from the presence of promotional signals that the price is lower than usual or from the absence of such signals that the price has increased.

- *Deal spotting*: this form of price knowledge consists of noticing that a price is within or outside the normal range of previous prices, by accessing the *analogue magnitude* code. (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.

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. Indeed, 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 is a function of the level of promotional activity in the category, i.e. how often a price changes. Price range is a measure of the difference in price from the lowest to highest priced item in the category. Category clutter is measured as the number of references 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. When the price of a product changes often over time, it should be more difficult to remember any given price from its price history (Johnson, 1994). However, if a product is often promoted on price, its price range should be salient (Kalyanaram and Little, 1994). Therefore, our prediction for price volatility depends on the measure considered. Price recall (H_{1a}) and recognition (H_{1b}) should be negatively affected by volatility, while deal spotting (H_{1c}) should benefit from the communication on price promotions (see appendix A for the formal hypotheses). 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. When the price range is large, the product category will be associated in memory with a large number of different prices, and this category therefore becomes an unreliable cue to access the price of any given brand in memory (Anderson and Bjork 1994). This should lead to a higher level of confusion about the actual price of any given brand. Category clutter is hypothesized to have a similar effect, because confusion again becomes more likely when the number of references 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 engaging in the most 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 research (H_{4a-c}) or across-store price research (H_{5a-c}) 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-c}).

Price Knowledge Survey

Summary of Study Features

As mentioned before, when designing our study, we tried to come as close as possible to the notion of reference price (without limiting ourselves to recall measures):

- Our questions were brand specific and for brands that had been purchased in the past by the respondents (unlike, for instance, in Urbany and Dickson 1991).
- The questions tapped into long-term memory and excluded the possibility that price was retrieved from short-term memory (unlike Dickson and Sawyer 1990 and their replications).
- To the extent that it was possible, we matched product and brand cues used in the questioning to those present in the choice context (e.g., through the use of actual photographs of the products tested).

- The questioning took place during a purchase occasion to maximize the number of contextual cues and to ensure the presence of normally available shopping knowledge.

Product Category Selection

The product categories for the survey were chosen so that they would represent high or low (but not medium) levels of each of 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 price volatility over time, average price range and number of references for about three hundred product categories. In a second step, we ran three store checks in the store where our survey took place over a period of 2 months to determine the eight most representative product categories(see Figure 1).

==========

Insert Figure 1 About Here

==========

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 supermarket, 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 compared 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. The interview was simply introduced as part of a study on consumer products. The introduction also mentioned that the respondent 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 shop 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 way of knowing what the actual prices are at the store they patronize on a regular basis. 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).

As already mentioned, our study 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 references sold by the store in the eight product selected categories. These were reproduced in an interview folder in alphabetical order

by brand, and, at a second level, by specific reference. Respondents were only interviewed on product categories they usually buy at the store. After a respondent had indicated which product reference s/he usually purchases, s/he had to give an estimate of its normal non-promotional price. Respondents could give the price they recalled in the unit they preferred: by item, by pack, by weight, or by volume. They also had to indicate whether they were sure, rather sure, or unsure about their estimate.

Next, a series of hypothetical prices for the product were presented in the interview folder, one at a time, and the respondent had to tell 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. 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. All price presentations and questions were presented orally, as well as in written form along with the different answer options in order to maximize comprehension of the task.

Results

In this section we proceed in two steps. First we report the overall results about the varieties of price knowledge of French consumers. After a discussion of the individual measures, we examine to which extent they 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 products selected were mineral water and yogurt, and one hundred consumers were interviewed for each. 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. From these data we can infer 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.

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

Figure 2 illustrates the respective roles of the difference in countries (French versus American) and in memory measures (short versus long-term) in explaining the difference between our study and Dickson and Sawyer. Both explanations clearly play a role. The fact that the slopes of the short and long-term curves in our study are not different ($p > 0.4$) indicates that the distribution of recall errors for short and long-term is almost similar, when compared against the respective perfect recall rates.

==========

Insert Figure 2 About Here

==========

Our recall data also give important indications of how people keep reference prices in memory. We noticed that a large majority of the recall estimates were in round numbers, while many of the actual prices contain decimals. In table 3 we categorized price estimates and actual prices by their price endings (X.00 for whole francs, X.X0 for prices in tenths of francs, and X.X5 for prices ending with 5 centimes). A comparison of the first two rows in this table shows that, in response to a recall question, most people do not respond with plausible prices, even if these are wrong, but instead give rounded estimates. This suggests that, when accessed in the form of retrieval, reference prices are not based on a representation of the actual price.

The recalled prices in the first row of Table 3 mix respondents with very different levels of price knowledge. Our previous finding then raises the question of whether respondents who demonstrate better price recall use a different storage/retrieval strategy than the less accurate respondents. One could expect that the former use rounding less frequently. A comparison of

the first with the third row of the table answers this question. This third row categorizes recall responses that were within 5% of the actual price. Apparently, more knowledgeable consumers do not use a different method, with less rounding, to store or retrieve price data (χ^2 =3.47, p=0.32). They use the same rounding strategies as everyone else and their knowledge advantage is therefore not based on using more precise numbers, but on simply using "better" numbers.

Table 3. Recalled and actual price endings

	X .00	*X.X0*	*X.X5*	*Other*
Recalled prices N=1166	77%	16%	3%	0.5%
Actual prices N=199	2%	33%	42%	23%
Recalled prices within 5% of actual N= 254	80%	15%	5%	0.3%

If most recall responses are influenced by a price rounding strategy, one should judge recall performance not by comparing it to absolute accuracy, but instead to a benchmark that also reflects this rounding strategy. We developed two benchmarks (Schindler and Kirby 1997; Stiving and Winer 1997). A first benchmark assumes that respondents round prices to the nearest price in Francs. The second benchmark takes the assumption that because of left to right reading, people remember the first digits better and round prices down. It turns out that when comparing to the nearest rounded price, 15.1% of the recall responses are accurate (up from 2.1% without rounding). Compared to floor rounding, accurate performance is at 15.2%. Taking price rounding into account clearly gives a much-improved picture of recall accuracy.

Another way to examine the impact of price rounding on recall is to consider what recall performance would be for someone who perfectly registers prices but always uses a rounding strategy (in either processing for storage or in retrieval). Across our sample, we find that using this strategy would have categorized only 5% of the responses as accurate, irrespective of the rounding strategy. Rounding to the nearest Franc would have brought 90% of the responses within the 5% error range, while floor rounding only would have brought 43% within that range. Again, when accounting for the role of price rounding, price recall is much more reliable than what the raw data suggest.

Our final analysis of the recall responses 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 unit, with the remainder stated by package. A *t*-test on absolute percent differences (abs[recall estimate - actual price]/actual price] shows that recall responses do not differ in accuracy when given per item or per package (per unit = 22, per pack = 26; $t_{512} = 1.76$, p = 0.08).

Price Recognition

As mentioned before, it may be premature to draw conclusions for retail price policy by just looking at the results for price recall. The other dimensions of price knowledge have also to be taken into consideration. Recognition of the actual price can be considered as the next step down on the price knowledge ladder. Table 4 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 took the hypothesis that a certain percentage of the respondents, X, recognize the actual price

with perfect confidence, while the complement uses pure guessing (Morrison 1981). Through random chance, one third of the guessers will pick the right answer. Thus, the number of correct responses is equal to X + (1 - X) / 3. In our case, the percentage of correct response is 42.2%; solving for X we have a percentage of correct non-guessers of 13.3%.

Table 4. Price Recognition

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

N = 1186

Comparing recognition with recall results gives 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 estimate that falls 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 reference price when they have to make a recognition judgment, even if this recalled price is fairly accurate. Put differently, if these consumers use a pictorial representation of the price to make their recognition judgment, this is apparently a less effective strategy than if they had used the spontaneously recalled price as a judgment basis.

Deal Spotting

Recall and recognition are useful forms of knowledge for the price-sensitive shopper 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 simply have a sense of which prices are attractive or unattractive. The response patterns to the four prices that were presented can be analyzed in different ways.

We start by dividing our responses into three groups by just considering the reactions to extreme changes of 20%. A first group is composed of consumers who are fairly knowledgeable about price magnitudes and respond positively to a large discount and negatively to a large increase in price. In our case this corresponds to a [good, —, —, bad] response to a price series of [-20%, -5%, +5%, +20%], or a [bad, —, —, good] response to a price series of [+20%, +5%, -5%, -20%]. Almost 36% of the responses 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% of our observations in this case. The remaining responses, 50%, reflect some intermediate form of deal spotting and constitute our third group. Figure 3 shows the breakdown on respondents among the three groups. Overall, the really clueless segment of respondents is rather small. Most of our sample is able to engage in deal spotting to some degree.

==========

Insert Figure 3 About Here

==========

The three basic groups of responses identified here can be further subdivided by considering all four responses. A retailer should obviously focus most on the knowledgeable shopper segment, that is, consumers with response pattern [good, —, —, bad] to an increasing price series (or the reverse to a decreasing series). Based on price knowledge alone, a knowledgeable consumer may have two reactions to a price that is five percent lower than normal: it can look like a good price or like a normal price. Similarly, a five percent higher price may seem like a high or a normal price. The answer will probably depend on the price sensitivity of the respondent. The response pattern that signals most price sensitivity is [good, good, bad, bad], and that with the lowest price sensitivity [good, normal, normal, bad]. Intermediate sensitivity will be revealed in the [good, good, normal, bad] and [good, normal, bad, bad] response patterns. A reverse response pattern should be obtained in each case to a decreasing price series. Table 5 illustrates how the knowledgeable responses can be classified on the price sensitivity dimension.

Table 5. Classification of knowledgeable responses on the deal spotting questions.

	Highest price sensitivity	Intermediate price sensitivity	Lowest price sensitivity
Calculated on knowledgeable responses[3]	4.5%	62%	26%
Calculated on total sample	2.0%	22%	9%

This table indicates that there are few consumers sensitive enough to price changes to react to a 5% price increase or decrease.

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 4). The most complete form of knowledge is the basis for accurate price *recall*. A second form permits accurate *recognition* of the actual price. A third form helps one to *react* correctly to prices that significantly depart from the usual price. The fourth form permits some level of sensitivity to *large* price differences but is not perfect. Finally, the fifth form indicates that even this level of sensitivity is *not* present.

The percentage of responses for the first three forms of knowledge seems to confirm that they are incremental steps on the price knowledge ladder. To check whether this is indeed the case we ran a Guttman scalogram analysis. In a Guttman scale, the items (memory measures in our case) are of increasing difficulty, which permits to position each respondent on an overall incremental scale based on all the items combined. To verify whether our memory measures are incremental, we calculated the coefficient of reproducibility 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 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 price recognition, deal spotting and price recall are of increasing difficulty, the Guttman scalogram analysis shows that instead, they are independent. 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).

==========

Insert Figure 4 About Here

==========

Drivers of Price Knowledge

We tested our hypotheses on price knowledge drivers by running logit regressions on our three measures of price knowledge (price recall, price recognition, and deal spotting). We used the knowledge drivers as independent variables. To study price recalled, we used as binary dependent variable whether responses were within a 5% range of accuracy or not[4]. For recognition, we look at whether the correct response was chosen or not (at an individual level, we cannot correct for guessing) and for deal spotting, we consider the difference between the 36% of responses that recognize good and bad deals versus the other responses. We also ran an analysis in which we coded, again with a binary variable, whether there was no deal spotting at all (14.1%) or some form of deal spotting. We expect answers of price oblivious shoppers to be completely random and that therefore none of the independent variables will be significant.

To avoid multi-colinearity problems in our analyses, we ran a factor analysis on the questionnaire items measuring the consumer-level knowledge drivers of hypotheses 4 to 6 (see Appendix B for details). This yielded the three factors we expected: (1) in-store price information search, (2) across-store price information search, and (3) shopping trip budget. We used 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 had access to national panel data with the average 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 prices, frequent switching would indicate a frequent processing of price information. On the other hand, frequent switching may also mean that consumers were exposed to many different prices, and thus it would be harder for them to remember any price accurately. Regarding the other variables, one would expect consumers to more accurately remember prices for products they purchased frequently, recently, and for product categories that are more salient (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 6.

The results shown in Table 6 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 is better for categories with more price volatility (H_{1a}) and is worse for categories with a larger price range (H_{2a}) and more references (H_{3a}). All these results are repeated for price recognition (H_{1b}-H_{3b}). This confirms our hypotheses for recognition, except for price volatility where we had predicted a negative effect (H_{1b}). We are not as successful in explaining deal spotting (H_{1c}-H_{3c}) from the category drivers, as we are with price recall and price recognition, but the direction of the effects is nevertheless consistent with our hypotheses, and the effect for category clutter is significant.

Table 6. 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**	-	-0.91*	-	-1.78*	-
Volatility **(H_1)**	1.17**	3.21	0.49**	1.63	0.35	1.42	0.05	1.05
Range **(H_2)**	-0.89**	0.41	-0.39*	0.68	-0.22	0.81	0.04	1.04
Clutter **(H_3)**	-1.32**	0.27	-0.62**	0.54	-0.41*	0.66	0.18	1.20
Frequency	0.09**	1.09	0.02*	1.02	0.03**	1.04	-0.02	0.98
Loyalty	0.08**	1.09	0.03**	1.03	0.008	1.01	0.01	1.01
Recency	0.04	1.04	0.01	1.01	-0.07	0.93	-0.12	0.89
Salience	0.15	1.17	-0.05	0.95	-0.03	0.97	0.16	1.17
In-store search **(H_4)** (Factor 1***)	-0.15*	0.86	-0.03	0.97	-0.17**	0.84	0.07	1.07
Across-store search **(H_5)** (Factor 2***)	0.09	1.10	-0.09	0.91	-0.04	0.96	-0.05	0.95
Shopping budget **(H_6)** (Factor 3)	0.04	1.04	-0.02	0.98	0.04	1.04	0.02	1.02
Chi-Square	p = 0.0001		p = 0.02		p = 0.0008		p = 0.65	

* Significant at p=0.05
** Significant at p =0.01
*** Because of reverse coding of the items of Factors 1 and 2, higher values signify less price search

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. In contrast, recency and salience turn out to be insignificant 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 consumers do hold reference prices in memory. Further, these reference prices are not an accurate (pictorial) representation of the last price seen for most people, but rather, seem to be generated through the processing and recoding of past price information. More research is needed to determine what the processing amounts to, but there is definitely some information loss through the rounding of prices. The amount of information lost

is large enough to make it difficult for consumers to recall or recognize actual prices (<15% of success), but 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 an 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). Interestingly enough, consumers who have a more precise knowledge of prices use the same rounding procedure as the consumers who have a less precise knowledge. 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.

Our analysis of the drivers of price knowledge indicates that frequent promotions make 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 an increased brand choice 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 these findings are related to product category, which means that the findings can be directly applied by retailers for their pricing strategies. This is not the case for shopper characteristics where retailers first have to measure 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. Without further research, this fact can be tentatively explained by the increase in task complexity for cherry-pickers. They now 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 the reference price. 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 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 would still be able to notice the promotional activity. In short, 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, Drèze, and Purk (1995), our findings are relevant to the work on inter-store competition. The important points here are that we found that consumers are not proficient at remembering regular prices, but are 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. 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).

Limitations

Before concluding this paper, we must be cognizant of the shortcomings of the methodology we used. Our basic instrument is a survey of consumers' stated price knowledge. This instrument is fraught with problems such as random responses, order effects, or stated versus revealed knowledge. We chose this instrument because we wanted to move from an *inferred* to a *direct* measure of reference price. We could have increased the validity of the results by asking only one price knowledge question (recall, recognition, or deal spotting) per respondent as that would have eliminated any risk of carry over from one answer to the other. Unfortunately, this would have rendered the survey prohibitively expensive to us. Further, we were limited in how we could rotate the questions to avoid order effects. We had to ask the recall question first because it would not make sense to show people three prices and ask them which one is the correct one, and then ask them a price recall question. Again, this may have induced carry-over effects. Finally, we restricted ourselves to consumers who made most of their shopping trips at the participating retailer. We did so to increase the validity of our comparison of the reference prices to the actual store price. 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 reference prices for the separate stores. We will leave this to future research. Another issue for later research is the necessity of making a better distinction between storage and retrieval effects. When we, for instance, observed rounding in recall, the most plausible explanation was that during the encoding and updating of price information, consumers simplified their task by rounding off the cents portion of a price. It is possible, however, that consumers store complete prices in memory but do not retrieve decimals when making purchase decisions

Conclusion

We showed in this study that reference prices do exist and that they are a long-term memory construct. The accuracy of consumers' reference prices has been shown to be dependent on both the environment (e.g., category clutter, promotion activity) and the idiosyncrasies of consumers (e.g., deal proneness, cherry-picking behavior). These findings are important to managers who have to make pricing decisions in a competitive multi-product environment. They also 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). By looking at different types of price knowledge (recall, recognition, and deal spotting), we show that although consumers do not possess an accurate knowledge of price, they possess a working knowledge of process that is generally inaccurate, but accurate enough to make good purchase decisions. Consumers may not know if the price is right, but most know if the price is wrong.

References

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

Kalwani, Manohar U., Chi Kin Yim, Heikki J. Rinne, and Yoshi Sujita (1990), "A Price Expectation Model of Consumer Brand Choice," *Journal of Marketing Research*, 27 (August), 251-262.

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.

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.

Monroe, Kent , 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.

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

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.

Schindler, Robert M. and Kirby, Patrick N. (1997), "Patterns of Rightmost Digits Used in Advertised Prices: Implications for Nine-Ending Effects," *Journal of Consumer Research*, 24 (September), 192-201.

Stiving, Mark and Russell S. Winer (1997), "An Empirical Analysis of Price Endings with Scanner Data," *Journal of Consumer Research*, 24 (June), 57-67.

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

Urbany, Joel E., Peter R. Dickson 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_{2a} The ability of a consumer to *recall* the price of an item is negatively affected by the range of price in the category in which the product belongs.

H_{2b} The ability of a consumer to *accurately recognize* the price of an item is negatively affected by the range of price in the category in which the product belongs.

H_{2c} The ability of a consumer to *spot a good or bad deal* on a product is negatively affected by the range of price in the category in which the product belongs.

H_{3a} The ability of a consumer to *recall* the price of an item is negatively affected by the level of clutter in the category in which the product belongs.

H_{3b} The ability of a consumer to *accurately recognize* the price of an item is negatively affected by the level of clutter in the category to which the product belongs.

H_{3c} The ability of a consumer to *spot a good or bad deal* on a product is negatively affected by the level of clutter in the category in which the product belongs.

H_{4a} Consumers engaging in in-store price research will be better at *recalling* the price of an item than those who don't.

H_{4b} Consumers engaging in in-store price research will be better at *accurately recognizing* the price of an item than those who don't.

H_{4c} Consumers engaging in in-store price research will be better at *spotting a good or bad deal* on a product than those who don't.

H_{5a} Consumers engaging in across-store price search will be better at *recalling* the price of an item than those who don't.

H_{5b} Consumers engaging in across-store price search will be better at *accurately recognizing* the price of an item than those who don't.

H_{5c} Consumers engaging in across-store price search will be better at *spotting a good or bad deal* on a product than those who don't.

H_{6a} Consumers who have large average basket sizes will be better at *recalling* the price of an item than those who don't.

H_{6b} Consumers who have large average basket sizes will be better at *accurately recognizing* the price of an item than those who don't.

H_{6c} Consumers who have large average basket sizes will be better at *spotting a good or bad deal* on a product than those who don't.

Appendix B: 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 take advantage of price promotions?	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 C: Summary of Hypothesis-Testing Results

	Effect on Price Recall	*Effect on Price Recognition*	*Effect on Deal Spotting*
H1: Category Volatility	*Support*	*Reverse*	No Support
H2: Price Range	*Support*	*Support*	No Support
H3: Category Clutter	*Support*	*Support*	*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

Figure 1: Category Selection

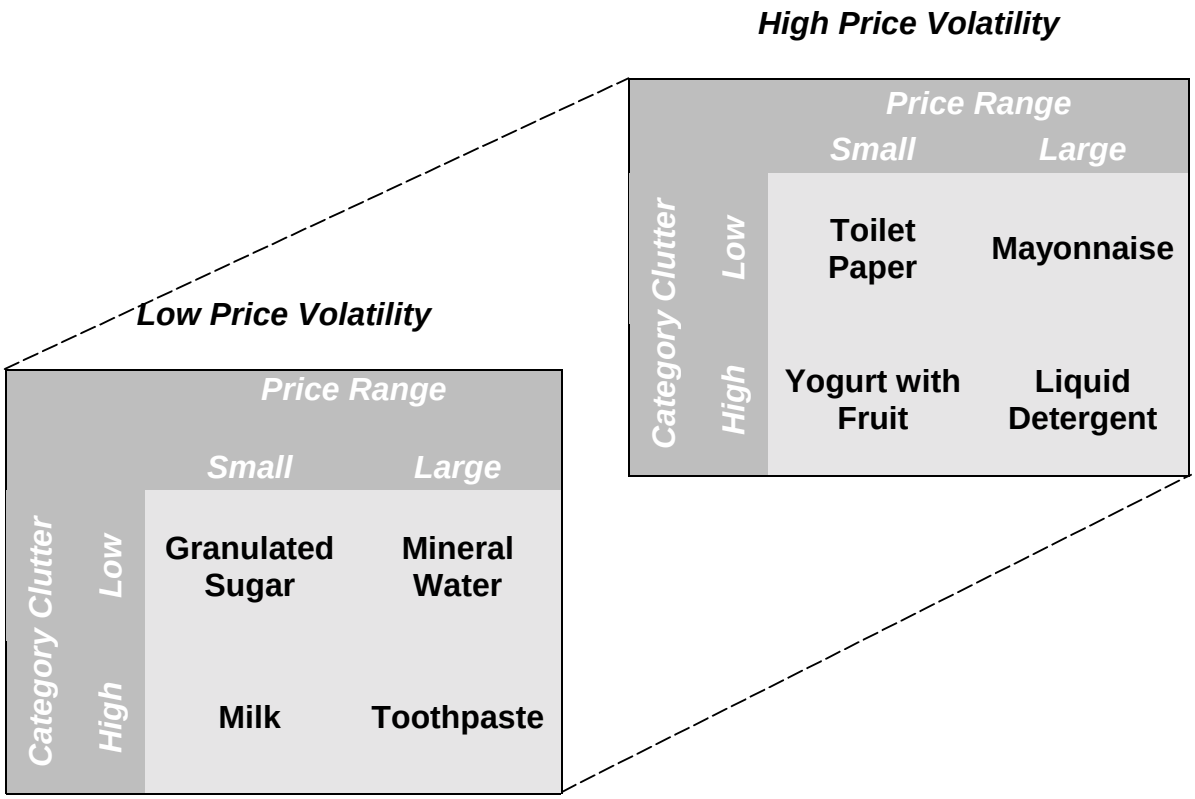


Figure 2: Comparison of Price Recall Accuracy

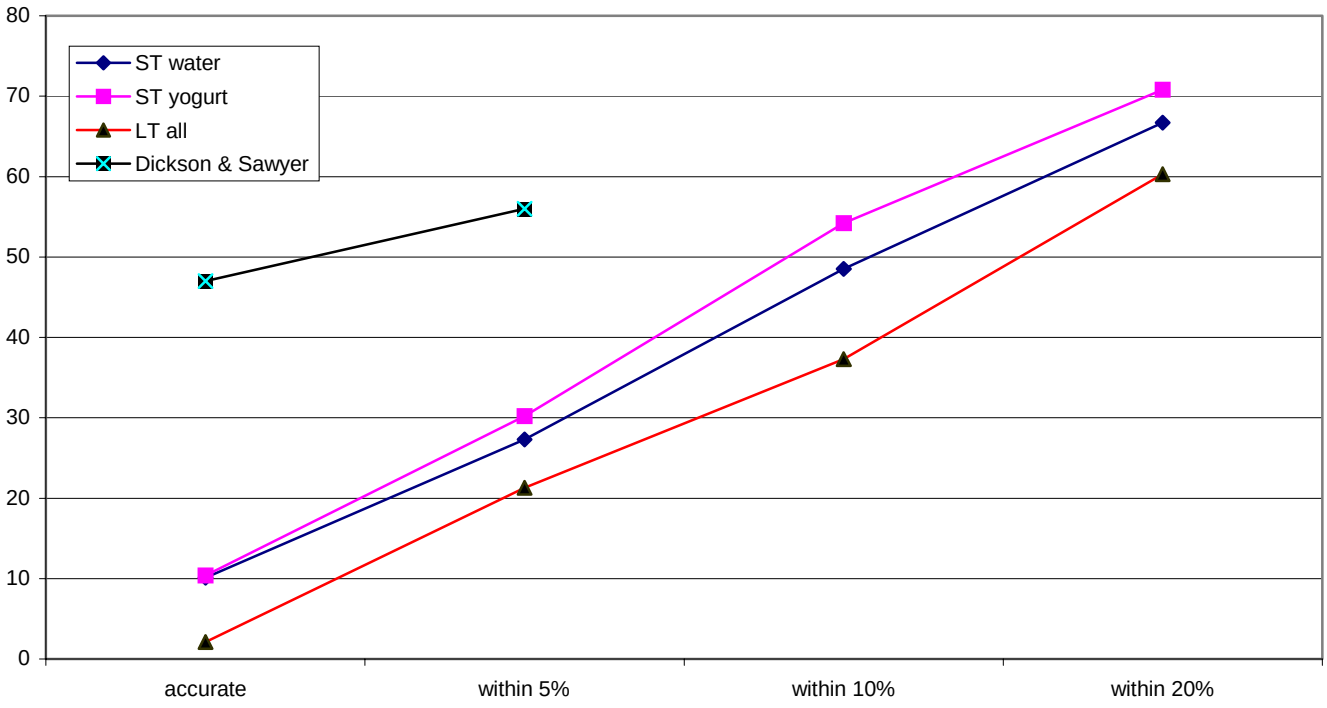


Figure 3: Deal Recognition Breakdown

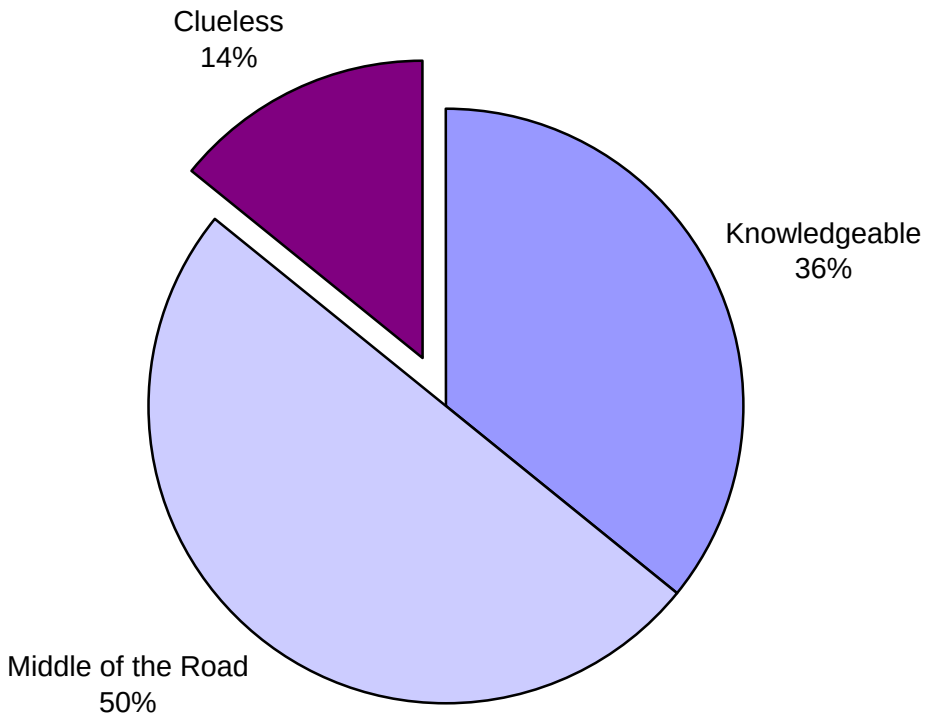
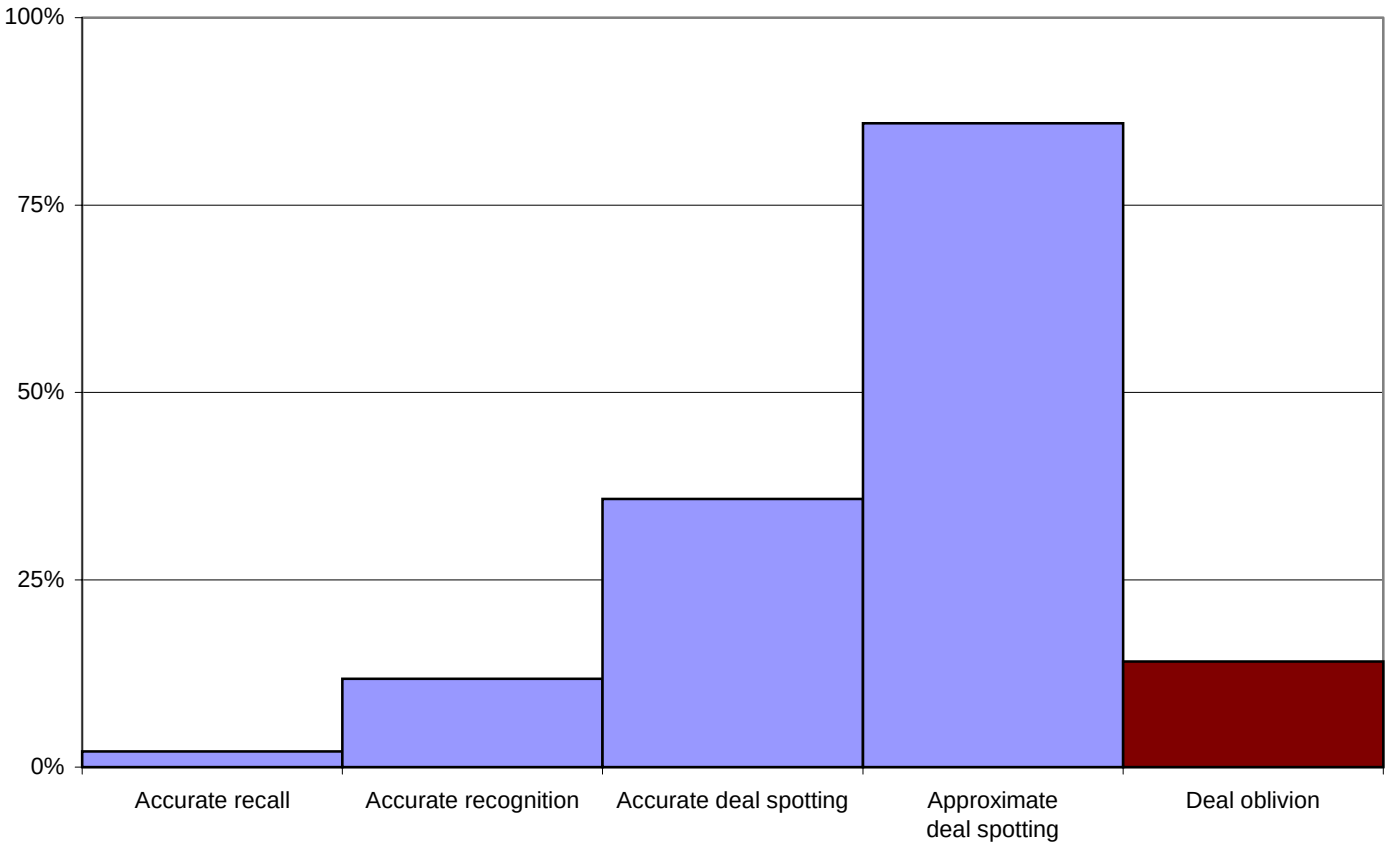


Figure 4: Levels of Price Knowledge



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] Percentages do not add-up to 100% as some respondents were not transitive (e.g., some respondents answered [Good, OK, Good, Bad] to prices of [-20%, -5%, +5%, +20%].

[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.