Consumers’ Immediate Memory for Prices

MARC VANHUELE
GILLES LAURENT
XAVIER DRÈZE*
*Marc Vanhuele is Associate Professor of Marketing at the HEC School of Management, Paris (GREGHEC), Marketing Department, 78351 Jouy-en-Josas, France (vanhuele@hec.fr). Gilles Laurent is the Carrefour Professor of Marketing at the HEC School of Management, Paris (GREGHEC), Marketing Department, 78351 Jouy-en-Josas, France (laurent@hec.fr). Xavier Drèze is Assistant Professor of Marketing, the Wharton School of Business, University of Pennsylvania, Philadelphia 19104-6340 (xdreze@wharton.upenn.edu). The authors thank Zsofia Kenesei and Sandor Czellar for their help with the data collection and Wes Hutchinson, David Luna, and Eric Yorkston for their comments on a previous draft.
Abstract

We examine the cognitive mechanics involved in keeping prices in short-term memory for subsequent recall. Consumers code and store prices verbally, visually, and in terms of their magnitude. The encoding used influences immediate recall performance. The memorability of prices depends on their verbal length, usualness, and on the other prices to be memorized simultaneously. The performance of consumers is affected by their auditory and visual recall ability, their pronunciation speed, and price abbreviation habits. Overall, consumers recall prices better than suggested by previous digit span studies with simple numbers.
Consumers make price-based decisions daily. They routinely observe and compare prices and make purchase decisions based on posted prices. Their accumulated price knowledge helps them determine whether a price they observe in a store, on a flyer, or in an advertisement is high or low. However, research has shown that consumers’ price knowledge with regard to packaged goods tends to be lower than what one might expect intuitively from the prevalence of customer–price interactions (Vanhuele and Drèze 2002). A couple of seconds after they pick up a product from a shelf, only 47% to 55% of consumers are able to recall its correct price (Dickson and Sawyer 1990; Le Boutillier, Le Boutillier, and Neslin 1994; Wakefield and Inman 1993). This may be due to a lack of attention and motivation by the consumer or to distractions in the store environment, but part of the explanation may also lie in the cognitive challenges of storing pairs of prices and product identifiers in memory and retrieving them.

Little is known about the cognitive processes involved in the encoding, storage, and retrieval of prices. Although conceptual models of the grocery shopper’s treatment of price information have been proposed (e.g., Dickson and Sawyer 1990; Monroe and Lee 1999), past research has focused on the output accuracy of the process. The cognitive processes themselves have not been examined. To advance our understanding of the price memory phenomena observed in the marketplace, we propose a detailed analysis of these cognitive processes. The logical starting point is the encoding phase.

Our objective is to identify an appropriate conceptual framework that relies on cognitive psychology and to verify that this framework effectively applies to price memory tasks that consumers typically perform. To do this, we study the cognitive processes involved in price encoding with realistic stimuli, in a setting relevant for typical consumer price interactions. We consider a consumer in an “optimal learning situation:” motivated to learn the prices of specific
products, placed in ideal conditions, concentrating on the learning task. We examine the cognitive mechanics involved in the storage and short-term recall of prices when a paired product serves as the recall cue. We are particularly interested in the capacity limits of working memory and the effects of these limits for short-term price recall. How many prices, and of what level of complexity, can consumers handle? For which types of prices are errors more likely? How do consumers differ in their capacity to retain prices in their memory for a couple of seconds?

To develop our conceptual framework we combined two fields of research in cognitive psychology: numerical cognition and the architecture of working memory. In the field of numerical cognition, a sub-domain of cognitive psychology, consensus is emerging about the way the human cognitive system deals with numerical information through a dedicated cognitive subsystem (Ashcraft 1992; Dehaene 1992, 1997). If prices are treated as numbers, memory performance for prices should reflect the architecture and processing characteristics of this cognitive system. Beyond its theoretical interest, this proposition also has practical importance in that it implies that specific prices will be more or less likely to be recalled accurately according to the way they are treated by the numerical cognitive subsystem. Recall errors therefore would contain a systematic component rather than being purely random.

The second theoretical basis for our work pertains to the architecture of working memory (Baddeley 1992, 2001). In a classic article, Miller (1956) proposed that the capacity of short-term memory has a fixed limit of approximately seven “chunks” or units of information. Baddeley shows, however, that immediate memory span is not constant in the number of units but rather depends on the phonological length of the words to be recalled. Many studies have confirmed Baddeley’s initial demonstrations that a person’s memory span for words equals the
number of words that person can read aloud in 1.5 to 2 seconds (Baddeley 2001). Using Baddeley’s work as a starting point, we examine the capacity limitations of working memory for pricing information.

Past research on numerical cognition and working memory cannot be applied directly to price knowledge because of the specific stimuli used in psychological experiments—typically one-digit integer stimuli without contextual meaning. In contrast, the prices consumers typically encounter in the marketplace are much more complex. In the terminology of numerical cognition, in addition to the lexical dimension (i.e., the number “vocabulary”), the combination of multiple digits involves a syntactical dimension (i.e., how individual elements are combined; 5 and 7 take a different meaning in $57 than in $572, McCloskey and Macaruso 1995). In addition, we did not find any psychological studies on number stimuli that combine an integer and a fractional part, such as a combination of dollars and cents (though this topic has received considerable attention in marketing; cf., Thomas and Morwitz 2005). Finally, in a consumption context, a price needs to be remembered in a pair with the corresponding product.

Two main contributions emerge from our work. On the theoretical side, this is the first attempt to combine the dominant theories in numerical cognition and on working memory. These theories prove to be compatible with one another and, moreover, applicable to an information processing based analysis of price memory. In combination, they generate predictions about the differences in memorability of different prices. They also predict how individual consumers differ in immediate memory performance. As a second contribution, we test and confirm most of these predictions in an experimental context that is relevant to the way consumers often interact with price. We show that prices are encoded in multiple ways and that each form of encoding
affects the way prices are remembered. These contributions should provide the basis for future information processing driven research on price memory.

THEORETICAL BACKGROUND

Central to the domain of numerical cognition is the question of how numbers are represented in the cognitive system and what roles the formats of this representation play in numerical processing (Ashcraft 1992; McCloskey and Macaruso 1995). In a synthesis of the essential research findings, Dehaene (1992) develops a triple-code model in which he proposes that numbers can be mentally represented and manipulated in three different forms, depending on the task at hand. The visual Arabic code represents numbers on a spatial visual medium on the basis of their written form in Arabic numerals (e.g., 35). The auditory verbal code is generated through a phonological representation in which each number is represented by a sequence of phonemes (e.g., /thirty/ /five/). Finally, the analogue magnitude code represents numbers as approximate quantities on an internal dimension termed the “number line” (e.g., about 35, slightly less than 40, or somewhere between 30 and 40). A representation in the analogue magnitude code apparently is generated automatically whenever a number is represented in the verbal or visual code.

Baddeley (1992, 2001) uses the term “working memory” to refer to the subsystem of the brain that temporarily stores and manipulates information. He proposes dividing working memory into three components: (1) the central executive that controls attention, (2) the visual-spatial sketch pad for manipulating visual information, and (3) the phonological loop, which consists of a phonological store to hold speech-based information for a duration of 1.5 to 2
seconds and an articulatory control process that can maintain data within the phonological store through subvocal repetition. Through the loop component, “subjects can generally remember as many words as they can say in 2 seconds” (Baddeley 1992, p. 558). Memory traces in the phonological store decay in approximately 2 seconds unless they are refreshed by sub-vocal rehearsal. Memory span in immediate recall tests then varies according to the length of the words used as stimuli (Baddeley, Thomson, and Buchanan 1975), because fewer long than short words can be rehearsed to refresh the phonological trace (Baddeley, Lewis, and Vallar 1984). Baddeley’s theory accounts for the findings in an impressive number of studies on immediate recall (Baddeley 2001).

Hypotheses

Analyses of each of the three codes proposed in numerical cognition, in combination with Baddeley’s work on the phonological loop, leads us to a set of predictions about the conditions under which consumers are likely to make errors when they attempt to memorize prices.

*Auditory verbal code.* An intriguing implication of the memory span limitation of 1.5 to 2 seconds is that prices that take longer to pronounce are less likely to fit in the phonological loop and therefore are less likely to be recalled accurately. Recording and measuring the
pronunciation durations for each price for each respondent in our empirical studies would be too cumbersome, so we use the number of syllables as proxy variable for most of our work\(^1\). Hence:

**H1a:** The more syllables in a price, the less likely it will be recalled accurately.

Most research that links pronunciation duration to recall (e.g., Schweickert and Boruff 1986) focuses on the effect of word length and ignores differences between persons. We are, however, also interested in the effect of individual differences in pronunciation rates.

**H1b:** Consumers who speak more slowly are less likely to recall prices accurately.

Different people may have different verbal price coding habits. A price of 137 could be coded in full as one hundred thirty-seven (7 syllables), or shortened to hundred thirty-seven (6), one thirty-seven (5), or one three seven (4). This habit variation leads us to another hypothesis.

**H1c:** Consumers who habitually shorten the verbal coding of prices will have better recall.

---

*Visual Arabic code.* In many countries, price endings tend to follow common structures. In the United States, for example, the digits 0, 5, and 9 are overrepresented as the rightmost price digits (Schindler and Kirby 1997). We hypothesize that the visual Arabic code will be particularly sensitive to those prices that stand out because they do not respect such common endings. These unusual prices should attract visual attention and therefore be more likely to be encoded visually, in addition to the verbal encoding. This double encoding should enhance recall (Chincotta et al. 1999; Frick 1984) provided there is enough encoding time.

---

\(^1\) We verified in a preliminary study with five consumers of different educational backgrounds and ages that there is a sufficient correlation (.89) between the number of syllables in a price and its pronunciation duration measured with millisecond-level accuracy.
**H2a:** Visually unusual patterns of prices will be recalled better than will visually usual patterns.

In addition to demonstrating the presence of visual coding and its effect on recall, we want to study the extent to which individual differences in visual memorization ability affect accuracy. Consumers have different abilities and preferences for visual versus verbal information processing (Childers, Houston, and Heckler 1985). In order to further explore the effect predicted by H2a, we develop a measure of visual memory. It is natural to assume that price recall will be correlated with this ability. More importantly, we expect an interaction effect between usualness and visual memorization ability. If consumers’ recall of unusual prices is driven by visual processing, those with a good visual memory should perform better with unusual prices.

**H2b:** The recall advantage for visually unusual patterns of prices will be more pronounced for people with a better visual memory.

*Analogue magnitude code.* Numbers in the analogue magnitude code are converted to approximate quantities and become less precise in absolute terms as the numbers get higher. More precisely, researchers on numerical cognition believe the hypothetical internal number line for magnitudes is scaled such that error sizes increase in direct proportion to price levels, a property called “scalar variability.” This is demonstrated by Dehaene and Marques (2002) in a study where respondents had to choose the typical price of a category from a series of seven proposed prices, and by Marques and Dehaene (2004) for a free price estimation task. Extending this to immediate recall, if the analogue magnitude code is used, the following should be true:

**H3:** Higher prices will lead to larger absolute deviations between the recalled and the actual price, and the deviations will be proportional to the price level.
In the remainder of this paper, we will test these hypotheses and examine how each of the three codes is involved in the encoding of prices. We present three studies that share a common research approach inspired by in-store price comparison making. Study 1 focuses on the phonological loop constraint. Study 2 examines the effect of price usualness and the role of visual coding. Study 3 zooms in on individual differences and tests for the presence of magnitude coding.

**STUDY 1: VERBAL CODING**

Overall Research Approach

All our experiments share a common approach. Because we ultimately want to understand the short-term recall challenges for consumers in the marketplace, we used in-store price comparisons as the inspiration for the experimental task. The different items considered by a consumer, while on the same shelf, are typically not located contiguously. As a consequence, if consumers want to compare their prices, they must store prices and product identifiers in short-term memory (consumers typically do not write down prices in the supermarket). Memory researchers refer to this task as “paired-associate learning.” To ensure that participants stored different price-product associations and not just a sequence of prices, we presented the product cues in the recall test one by one in random order.

To keep the task as realistic and engaging as possible, we used as stimuli pictures of real products that exist in marketplace of the country in which we ran each study. We took the actual prices as starting points to avoid confusing participants with unrealistic prices levels and then
made modifications to operationalize the price length manipulation. While our decision to use pictures of real products standardizes the stimuli, a drawback is that we could not control the memory space that the identification of the product itself requires. That is, we do not know whether individual respondents memorized the names of the products or took a mental picture of them; any variation in recall performance due to the identification of the product is therefore part of the experimental noise.

The instructions explained that the study had been designed to examine the capacity of consumers to memorize prices. Two or three products and their prices would be presented together on a screen. Participants then would view the products again one by one, in random order, and be asked to recall the prices. They were told that they always had to give an answer. Before the actual experiment started, they watched a demonstration of a complete trial (with music CDs, a category not used in the tests), during which the instructions were repeated.

Design

The main experimental variables of interest in this study are the price length of the target product in number of syllables (3–8) and the price length of the other products on the study screen (3–8 syllables each). Recall performance might also be influenced by the number of products per study screen (2 or 3), the presentation order (from left to right, 1, 2, or 3), and the question order (1, 2, or 3), which affects the time the consumer has to rehearse an item before recalling it. Therefore, we introduce these factors as control variables.

We used three product categories: (1) candy, with prices that comprised one dollar-level digit and two cents digits (e.g., $1.19); (2) DVDs with two dollar digits and two cents digits (e.g.,
$12.95); and (3) digital cameras with three or four dollar digits and zero cents digits (e.g., $345).

To ensure that the price stimuli were realistic, we started with the actual price of the products and then induced changes with minimal monetary impacts to attain the desired number of syllables. The resulting prices ranged from $1.05 to $3.95 for candy, from $12.15 to $72.26 for DVDs (and DVD collections), and from $120 to $2,390 for cameras. Product shots of each were downloaded from the Internet.

To counterbalance the presentation order and keep the number of syllables of a given price independent of that of the other price(s) on a particular study screen, we used the following procedure: For each product, we created a price we labeled “short” (S) and another we labeled “long” (L) (this distinction between “short” and “long” is only made to operationalize the counterbalancing in the experimental design; for the statistical analysis, we use as an explanatory variable the actual number of syllables). For study screens that displayed two products, we used the following four price/length combinations: SS, SL, LS, and LL. For three-product screens, we provided eight combinations from SSS to LLL. Each respondent viewed 12 screens, corresponding to all 12 price/length combinations for a given product category, which resulted in a total of 32 product–price pairs (four two-product screens, eight three-product screens). Our database contained 32 visuals for each product category, and for each visual we chose a “short” and a “long” price. For each respondent, we randomly assigned each visual to one of the two- or three-product screens and to one of the “short” or “long” price conditions.

To limit the overall duration of the task, we used a fractional factorial design in which we showed each respondent only two of the three product categories. Each respondent therefore viewed a total of 24 screens, in random order, with in total 64 product-price pairs to recall.
Procedure

Ninety-one U.S. undergraduate business students participated for course credit in this online experiment. The introductory instructions explained that the objective was to examine the capacity of consumers to hold prices in memory for a couple of seconds. Participants were told that they would see screens with two or three products and their prices and that they had to memorize the prices for immediate recall. They were informed that their answer had to be as accurate as possible. Because the experiment focuses on the role of the verbal length of a price, we instructed participants to read prices to themselves during the study phase “without speaking out loud” (this reflects the way they read prices in a store environment). We told them that they then had to rehearse these prices internally to keep them in memory.

Each trial started with a two-second announcement of the product category, followed by a nine-second presentation of the screen with the product–price pairs, a three-second masking of the prices (to clear the visual sensory buffer; see Breitmeyer and Ogmen 2000; Sperling 1960), a 2-second blank screen, and then the presentations in random order of each product for recall. Respondents had to type their answer for each product before being shown the next product. After typing the last answer for a trial, participants clicked a continue button to move to the next trial.

Analysis

We coded recall performance as a binary variable (correct or incorrect) and used logistic regression to test our hypotheses. Overall, 49.8% of the responses were correct. Prices with
more syllables are indeed less likely to be recalled. This is demonstrated (see Table 1) by the significant parameter for the number of syllables in the focal item ($\chi^2(1) = 95.60, p < .01$) and provides support for H1a. The effect size is considerable as each additional syllable in the focal item reduces the odds of recall by 20% (exponent of the parameter). Moreover, each additional syllable in the other item(s) leads to another reduction of 12% ($\chi^2(1) = 51.18, p < .01$). These effects operate in addition to the effect of the number of product-price pairs presented. Presenting three prices to remember, instead of two, almost triples the odds of a recall error. Each of these three effects captures part of the effect of the phonological loop constraint.

As for the control variables, the first recall question asked has a much higher chance of producing a correct answer than do either the second or the third ($\chi^2(1) = 26.69, p < .01$). This order effect probably results from the passage of time, the ensuing decay of the memory trace in short-term memory, and the increase of interference among prices. These effects therefore may be considered as additional reflections of the limitations of the phonological loop (see Cowan et al. 1992): a longer output time forces a participant to rehearse an item for a longer period of time and makes interferences and/or decay more likely. There is a significant effect of product category ($\chi^2(2) = 121.80, p < .01$), in that recall for cameras is best, followed by that for candy and then for DVDs. These recall differences may be attributed to various factors, including the number of digits, the presence or absence of cents, price levels, problems of visual identification, and recall of the product identification.
STUDY 2: VISUAL MEMORY

Our second study has two objectives. First, we test H2a and H2b regarding the effect of visual memory. Second, given our previous demonstration of the effect of the number of syllables of the price on price recall, we became interested in examining how consumers in countries with currencies with high face value deal with the challenges of price recall. Our first recall test used a relatively high-valued currency—the dollar—but in some other currencies the prices of similar products include many more digits. The Hungarian marketplace is ideal to examine this dimension. The Hungarian forint is pegged to the euro at an exchange rate of approximately 250 forint per euro (i.e., 202 forint per dollar at the time of the study). In addition, price endings in Hungary follow a restricted number of patterns, which enables us to manipulate the visual “usualness” of prices in the Arabic format, and therefore to test H2a and H2b. Hungarian prices do not use decimals and unit prices that are three digits long (e.g., for the candy product category used in our study) always end in a 5 or 9. Prices with five or six digits (e.g., digital cameras) always end in 90 or 00. We created “unusual” prices by using other endings (e.g., 223 or 84,560).

Design and Procedure

The experimental instructions were identical to those of Study 1. Participants were directed to use verbal coding and rehearse the prices verbally. Thus, if visual coding still plays a role, this study represents a strong demonstration of the role of visual memory. Due to practical
constraints participants could not be tested individually and were tested in groups, which precluded a full randomization of prices across word lengths and presentation and test orders. Participants viewed a PowerPoint slide show with the instructions, study, and test screens. The price usualness manipulation was between-subjects. One group only saw prices with usual endings, and the other saw a mix of prices with usual and unusual endings. We will subsequently refer to them as the usual and unusual groups respectively. Two product categories were used: candy and cameras. Within each category we used the staircase method in which the total number of syllables to be remembered on each subsequent screen increases (Baddeley, Thomson, and Buchanan 1975). The first trials for candy contained 8 syllables and those for cameras 17 syllables. The last trials had 21 and 42 syllables respectively. The first three screens for each category included two stimuli, and the final three had three. Each screen for the unusual group provided one unusual price with an equal number of syllables as the corresponding price for the usual group (e.g., 137 instead of 135—in Hungarian, 137 and 135 have the same number of syllables).

Scale for Visual Memory Capacity

In addition to testing for the main effect of price usualness, we also test its interaction with the participants’ visual memory capacity. To measure this capacity we developed a modified version of the Corsi block-tapping test, which had been designed to assess visual memory span (Berch, Krikorian, and Huha 1998). In the first part of the test, participants successively viewed six tables with four columns and four rows each, with some cells colored black and others in white. After seeing each table, respondents had to indicate which cells were
black. In the second part of the test, they were shown sequences of four or five tables, each with one black cell. Respondents were asked to replicate the sequence in which the black cells appeared by numbering the cells of an empty table. We checked the internal validity of our measure in a pilot test (Cronbach’s alpha = 0.71).

Analysis

Due to the design of the Hungarian study, there is a high correlation between the number of syllables of the target and other price(s) in a given trial (candy prices are three digits long; camera prices are six digits). In the logistic regression (see Table 2), we therefore used only one variable for syllable length to code the total syllable length of all prices in a given trial. We replicated the results of the U.S. study for syllable length, presentation order, and question order. We also found support for our hypothesis H2a that unusual prices have a recall advantage. Indeed, compared to the usual prices that all participants saw, the unusual prices seen by the unusual group had recall odds that were 69% higher ($\chi^2(1) = 7.94, p < .01$). However, overall there was no difference between the usual and unusual groups in recall performance. Probably the most plausible explanation for this pattern of results is that unusual prices draw attention to the detriment of usual prices.

We attribute the observed recall advantage for unusual prices to visual encoding that complements verbal encoding (H2b). To test this explanation, we used as an explanatory variable
our modified version of the Corsi test. Visual memory ability plays a significant positive role in overall memory performance ($\chi^2(1) = 8.92, p < .01$; note that our measure of visual memory ability is expressed as the number of errors). The interaction effect predicted by H2b implies that people with good visual memory ability should perform especially well with unusual prices. The results actually show the opposite effect ($\chi^2(1) = 3.87, p < .05$). A possible explanation is that the Corsi test does not measure visual memory ability as such but rather the ability to pick up on patterns. What differentiates usual from unusual prices is that they end in well-known patterns. Respondents scoring high on the Corsi test then also have better recall for usual prices.

STUDY 3: INDIVIDUAL DIFFERENCES IN RECALL STRATEGY

In our third study, we test H1b regarding speech rates, H1c about verbal coding habits, and H3 pertaining to magnitude coding. An additional objective is to examine to what extent immediate price recall is driven by the verbal, visual, and magnitude codes. For this purpose we introduce two new measures: one of verbal memory ability and a second of visual magnitude estimation capacity. Ninety-eight American students participated in return for course credit. The price recall part of the experiment was identical to Study 1 except that we did not give any instructions about how the prices were to be coded or rehearsed. Participants therefore used their personal “natural” encoding strategies. After the price recall part of the study, respondents had to answer five series of questions focused on individual differences. First, we measured the participants’ habitual way of pronouncing prices by giving them four 3- and 4-digit prices, asking them to read the prices to themselves, and then presenting them with alternative ways to

2 With thanks to an anonymous reviewer who suggested this explanation.
pronounce the price (e.g., 1290 as twelve ninety, twelve-hundred ninety, one thousand two hundred ninety). The participant also had the option of writing another pronunciation in words in a box labeled “other.” We counted the total number of syllables in each participant’s habitual pronunciation across the three prices. This was followed by the modified Corsi test with six fixed and six dynamic tables (Cronbach’s alpha = 0.83), an auditory recall test, a visual magnitude estimation test, and a measure of the participant’s speed of pronunciation.

To measure verbal memory ability, we used auditory recall to exclude the interference of visual memory. Our test was similar to the classical Wechsler (1958) test, which is widely applied in clinical diagnosis, but we used letters instead of numbers. The computer played sound files of someone reading strings of letters at a rate of one letter per second. In the first part, participants were to memorize the letters and simply enter them into the computer in the same order. In the second part, they had to enter the letters in reverse sequence. For each part, an ascending number of letters was presented (from 5 to 8), and there were two strings for each string size. We used the total number of correct answers to score participants (Cronbach’s alpha = 0.68).

To assess pronunciation rate, we measured the length of time the respondents needed to read to themselves two 1-page lists with 45 prices from the recall part of the test. The correlation between the two time measures across persons was 0.75.

By including three product categories with different price levels, we obtained a broad spread of prices, which enabled us to create a better test of the scalar variability than if we had used just one product category.

Because we did not find any established tests to estimate people’s ability to develop good magnitude estimates, we developed our own, in which we present people with pictures of a set of
identical objects that are too numerous to count during the presentation time. Because the
obtained estimates are not normally distributed, we rescale them by taking the logarithm of the
ratio of the estimate to the actual value before adding them and checking the reliability of our
scale (Cronbach’s alpha = 0.79). In the logistic regression, we classified respondents into three
groups based on the size and direction of their estimation errors: the 25% who give the largest
underestimates, the 25% who give the largest overestimates, and the rest, whom we use as a
reference category.

Analysis

In Table 3, we show the replication of the previous results for the price length effect and
the primacy effects for presentation and question order. The effect of speech rate (H1b) is
significant though small in size ($\chi^2(1) = 4.21, p < .05$). Each additional ten seconds needed to
read 45 prices corresponds to a 3% decrease in accuracy. Respondents who habitually use more
syllables to read a price present a lower recall performance (H1c, $\chi^2(1) = 4.45, p < .05$). The
effect size is almost identical to that of the word length effect; for each extra syllable in coding,
there is a 9% reduction in recall odds. By including the participants’ verbal, visual, and
magnitude coding abilities in our logistic regression, we establish how they explain individual
differences in recall performance.

Insert table 3 about here

Respondents’ performance on the auditory memory test is a strong positive predictor of
price recall performance. Visual memory ability also plays a significant positive role in price
recall. However, our visual magnitude estimation test does not have significant predictive power for recall. The parameter estimates suggest that over- and under-estimators suffer a lower recall level, but the effects are not significant.

Our final hypothesis, H3, is based on the assumption that some form of magnitude coding takes place, which should be reflected in recall error sizes that are proportional to the price level. To test this hypothesis, we ran a log-log regression (Figure 1) linking the standard deviation of the recalled prices to the level of these prices, as proposed by Dehaene and Marques (2002). Scalar variability predicts a slope of 1. With its value of 1.06, the slope is close to 1, although the standard error of the estimate (0.019) makes it significantly different from the predicted value. In conclusion, the pattern of results indicates the presence of magnitude coding.

In conclusion, the pattern of results indicates the presence of magnitude coding.

Insert figure 1 about here

ESTIMATING IMMEDIATE MEMORY CAPACITY

In Figure 2, we compare the average recall performance of our participants in the three studies by plotting recall as a function of the number of syllables of the price. Memory span usually is defined as the number of items for which the probability of correctly ordered recall in serial recall tasks is at least 50% (Schweickert and Boruff 1986). Using this criterion, the average
memory span in the U.S. studies is 13 syllables. In contrast, the performance of the Hungarian participants is startlingly high with a digit span of 24 syllables. To compare these studies on a more equal basis, we could base our estimate of the Hungarian digit span on just the “usual price” condition, but this only reduces their span to 19.5 syllables. How can we explain this difference?

One possibility is that Hungarians speak faster. Research on the memory span across different languages has suggested that languages that use more syllables on average per word tend to be spoken at a faster rate, possibly to attain “an optimum rate of information flow … to suit the human capacity for message transmission and reception” (Naveh-Benjamin and Ayres 1986, p. 749). Because we did not find any published work on speech rate or language “length” that included Hungarian participants, we compared the pronunciation speed between Hungarian and American participants for the prices used in the previous studies to test the speech rate explanation. We asked 10 people in each country to read sub-vocally three samples of the prices used in the experiments. The total pronunciation time was then divided by the number of syllables, which resulted in an average sub-vocal reading time per syllable of 0.198 seconds for the Hungarians and 0.279 seconds for the Americans. Thus, the Hungarians clearly are faster speakers, but the question remains whether speech rate explains the differences in immediate recall performance. If we multiply the sub-vocal reading time per syllable by the observed memory span, in terms of the number of syllables in the recall studies, we obtain a measure of memory span that has been corrected for differences in speech rate. For American speakers, memory span for prices is 3.72 seconds, whereas for Hungarian speakers, it is 3.86 seconds for usual prices, a difference that is not significant ($t(18) = -.563, p < .5$). In conclusion, we can explain the performance difference through the differences in pronunciation speed.
In addition to explaining part of the performance difference between Hungarian and American participants, our measures of pronunciation duration also highlight that the immediate recall of prices is far superior to that for single digits, as documented in memory span research, even though our recall task is more difficult for participants. In the typical memory span test, participants are exposed to a list of items that they must feed back immediately in the correct order. Our immediate recall procedure is more demanding because it includes more material to remember and, on average, must be retained in memory for a longer time (i.e., while the first recalled price is entered, the second, and possibly third, must be rehearsed). The memory span for single digits has been estimated at approximately seven (e.g., Schweickert and Boruff, Figure 2), which corresponds to 7.77 syllables in English (one for each digit except 7), in contrast to 13 or 24 syllables in our studies.

**SUMMARY OF MAIN CONTRIBUTIONS**

Consumer researchers have shown active interest in measuring the extent to which consumers can recall prices (see Monroe and Lee 1999 for a review). However, very little work has examined the memory processes that account for observed levels of performance. We focus on these cognitive processes by asking consumers to perform a realistic task (memorizing for a brief period the prices of two or three competing products in the same category, which is what shoppers need to do, in front of the shelf, in order to compare prices), in an optimal learning situation, and observing how their recall performance varies as a function of the price stimuli.

The combination of Dehaene’s triple-code model and Baddeley’s description of the architecture of working memory proves to be very productive for the study of immediate
memory of prices. Baddeley’s work describes the “hardware” components that operate at the encoding side for the three memory codes proposed by Dehaene. Overall, we find that variations in immediate recall performance reflect the cognitive mechanisms that researchers in numerical cognition and working memory have posited. Interestingly, our analysis in the context of pricing highlights the role of two factors that these researchers have not considered in the past: the structural features of numbers and the differences among individuals.

We find evidence that consumers code prices in all three codes—verbal, visual, and magnitude—and that different consumers have different tendencies or abilities to rely on each code. The overwhelming evidence of verbal coding reveals itself most clearly in the effect of the time needed to pronounce a price on its likelihood of recall. This effect takes three different forms in our studies. First, each extra syllable in a price decreases its chances of being recalled by 20%. Second, consumers can use shortcuts to reduce the number of syllables they must pronounce; in the most extreme form, they could ignore syntactic conventions and read out multi-digit prices as a sequence of single digits. Third, there is some cognitive cost of pronouncing prices slowly because it leads to the inefficient use of the phonological loop. In addition to these three effects, we observe that good auditory coders have better price recall performance.

We also find evidence of visual coding in two types of effects. First, we observe the attention effect that unusual prices generate. Second, we show that good visual coders perform better on the recall test. To detect the presence of magnitude coding, we use an indirect approach and demonstrate that recall errors are proportional to the price level, a prediction we derived from the hypothetical number line generated by magnitude coding, which gets compressed when numbers get higher and results in a degraded precision in coding.
Across our three experiments, we find that immediate recall performance is higher than predicted by digit span studies, even when we ignore that, in addition to a price, a product identification also had to be kept in memory. This level of performance most likely can be attributed to the combined use of the three types of coding. In addition, certain visual aspects of the price can provide additional cues that, again supported by long-term memory, help reconstruct the recalled price.

Because so little work has been done on price memory, we examined immediate or short-term recall as the first step. Our project has nevertheless also implications for long-term recall phenomena. Presence in the short-term store is a prerequisite for long-term storage and the latter form of storage should be a reflection of the way a price was represented in the short-term store. We hope that our conceptual framework and experimental approach can be the basis for future research in this direction.
REFERENCES


### TABLE 1
LOGISTIC REGRESSION ANALYSIS STUDY 1 (U.S.)

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>Wald</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of syllables of focal item</td>
<td>-.222</td>
<td>95.60*</td>
<td>.801</td>
</tr>
<tr>
<td>Number of syllables of other items</td>
<td>-.121</td>
<td>51.18*</td>
<td>.886</td>
</tr>
<tr>
<td>Number of items for trial†</td>
<td>1.062</td>
<td>93.01*</td>
<td>2.893</td>
</tr>
<tr>
<td>Presentation order†</td>
<td>3.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First</td>
<td>.134</td>
<td>2.77</td>
<td>1.144</td>
</tr>
<tr>
<td>Second</td>
<td>.017</td>
<td>.04</td>
<td>1.017</td>
</tr>
<tr>
<td>Question order†</td>
<td>33.04*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First</td>
<td>.418</td>
<td>26.69*</td>
<td>1.519</td>
</tr>
<tr>
<td>Second</td>
<td>.075</td>
<td>.74</td>
<td>1.078</td>
</tr>
<tr>
<td>Product category†</td>
<td>121.80*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Candy</td>
<td>-.265</td>
<td>14.46*</td>
<td>.767</td>
</tr>
<tr>
<td>DVD</td>
<td>-.772</td>
<td>120.38*</td>
<td>.462</td>
</tr>
<tr>
<td>Intercept</td>
<td>1.865</td>
<td>77.93*</td>
<td>6.457</td>
</tr>
</tbody>
</table>

N= 6248, Nagelkerke $R^2 = .22$

* Significant at 0.05 level
† The reference category is Camera, three items, third presentation, and third question order.
<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>Wald</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of syllables of items</td>
<td>-.131</td>
<td>418.17*</td>
<td>.877</td>
</tr>
<tr>
<td>Viewing only usual prices⁺</td>
<td>.128</td>
<td>2.12</td>
<td>1.136</td>
</tr>
<tr>
<td>Unusual focal pricesₓ</td>
<td>.526</td>
<td>7.94*</td>
<td>1.692</td>
</tr>
<tr>
<td>Presentation order</td>
<td></td>
<td>21.86*</td>
<td></td>
</tr>
<tr>
<td>First</td>
<td>.452</td>
<td>12.84*</td>
<td>1.571</td>
</tr>
<tr>
<td>Second</td>
<td>.054</td>
<td>.21</td>
<td>1.056</td>
</tr>
<tr>
<td>Question order</td>
<td></td>
<td>9.93*</td>
<td></td>
</tr>
<tr>
<td>First</td>
<td>.213</td>
<td>2.76</td>
<td>1.237</td>
</tr>
<tr>
<td>Second</td>
<td>.364</td>
<td>9.12*</td>
<td>1.438</td>
</tr>
<tr>
<td>Candy</td>
<td>.217</td>
<td>4.55*</td>
<td>1.243</td>
</tr>
<tr>
<td>Errors Corsi test</td>
<td>-.013</td>
<td>8.92*</td>
<td>.987</td>
</tr>
<tr>
<td>Corsi test × unusual prices</td>
<td>.020</td>
<td>3.87*</td>
<td>1.020</td>
</tr>
<tr>
<td>Intercept</td>
<td>2.049</td>
<td>47.00*</td>
<td>7.762</td>
</tr>
</tbody>
</table>

N = 4050, Nagelkerke R² = .40

* Significant at .05 level
⁺ The reference category is the respondent group that viewed a mix of usual and unusual prices.
ₓ The reference category is a focal usual price.
### TABLE 3
LOGISTIC REGRESSION ANALYSIS STUDY 3 (U.S.)

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>Wald</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of syllables of focal item</td>
<td>-.099</td>
<td>18.46*</td>
<td>.905</td>
</tr>
<tr>
<td>Number of syllables of other items</td>
<td>-.090</td>
<td>26.97*</td>
<td>.914</td>
</tr>
<tr>
<td>Number of items for trial</td>
<td>1.282</td>
<td>126.30*</td>
<td>3.603</td>
</tr>
<tr>
<td>Presentation order</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First</td>
<td>.333</td>
<td>20.22*</td>
<td>1.395</td>
</tr>
<tr>
<td>Second</td>
<td>.076</td>
<td>1.06</td>
<td>1.079</td>
</tr>
<tr>
<td>Question order</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First</td>
<td>.306</td>
<td>17.10*</td>
<td>1.358</td>
</tr>
<tr>
<td>Second</td>
<td>.141</td>
<td>3.63</td>
<td>1.152</td>
</tr>
<tr>
<td>Product category</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Candy</td>
<td>.042</td>
<td>.34</td>
<td>1.043</td>
</tr>
<tr>
<td>DVD</td>
<td>-.978</td>
<td>181.94*</td>
<td>.376</td>
</tr>
<tr>
<td>Errors Corsi test</td>
<td>-.080</td>
<td>17.61*</td>
<td>.923</td>
</tr>
<tr>
<td>Reading time for price tables</td>
<td>-.003</td>
<td>4.21*</td>
<td>.997</td>
</tr>
<tr>
<td>Syllable count with usual pronunciation</td>
<td>-.097</td>
<td>4.45*</td>
<td>.908</td>
</tr>
<tr>
<td>Performance auditory recall</td>
<td>.143</td>
<td>119.10*</td>
<td>1.153</td>
</tr>
<tr>
<td>Performance magnitude estimation</td>
<td></td>
<td>.13</td>
<td></td>
</tr>
<tr>
<td>Underestimators</td>
<td>-.010</td>
<td>.02</td>
<td>.990</td>
</tr>
<tr>
<td>Overestimators</td>
<td>-.029</td>
<td>.13</td>
<td>.971</td>
</tr>
<tr>
<td>Intercept</td>
<td>1.131</td>
<td>11.30*</td>
<td>3.098</td>
</tr>
</tbody>
</table>

N = 6278, Nagelkerke $R^2 = .26$

* Significant at .05 level
+ The reference category is a camera, three items, third presentation and third question order, and the middle group in the magnitude estimation.
FIGURE 1

SCALAR VARIABILITY IN IMMEDIATE PRICE RECALL

![Scatter plot showing variability in price recall for different categories: DVD, Candy, Camera. The x-axis represents LMEAN, and the y-axis represents LSTDEV. Different symbols are used to distinguish between categories.](image)
FIGURE 2
RECALL PERFORMANCE AS FUNCTION OF TOTAL NUMBER OF SYLLABLES TO BE REMEMBERED