Acta Psychologica 131 (2009) 221–234

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

A feature-based inference model of numerical estimation: The split-seed effect

Kyle B. Murray *, Norman R. Brown [1]

University of Alberta, Edmonton, AB, Canada T6G 2R6

ARTICLE INFO

Article history:
Received 5 December 2008
Received in revised form 28 April 2009
Accepted 21 May 2009

PsycINFO classification:
2340
3900

Keywords:
Cognitive psychology
Judgement and decision making
Feature-based inference
Numerical estimation
Consumer psychology
Pricing

ABSTRACT

Prior research has identified two modes of quantitative estimation: numerical retrieval and ordinal conversion. In this paper we introduce a third mode, which operates by a feature-based inference process. In contrast to prior research, the results of three experiments demonstrate that people estimate automobile prices by combining metric information associated with two critical features: product class and brand status. In addition, Experiments 2 and 3 demonstrated that when participants are *seeded* with the actual current base price of one of the to-be-estimated vehicles, they respond by revising the general metric and splitting the information carried by the seed between the two critical features. As a result, the degree of post-seeding revision is directly related to the number of these features that the seed and the transfer items have in common. The paper concludes with a general discussion of the practical and theoretical implications of our findings.

1. Introduction

The ability to generate numerical estimates is a critical component of our capacity to understand our physical and social environments. Such judgments are an important and pervasive part of our day-to-day lives. Previous research has examined numerical estimation in a wide variety of situations ranging from the nutritional value of fast foods (Wallbaum, 1997) and number of lifetime sexual partners (Brown & Sinclair, 1999) to insect toxicity (von Helversen & Rieskamp, 2008) and dates for personal and public events (Brown, 1990; Burt, 1992). The results indicate that in some domains people are relatively accurate (see, e.g., university tuitions, Lawson & Bhagat, 2002), while in other domains respondents can err by an order of magnitude or more (see, e.g., estimates of national populations, Brown & Siegler, 1993). Despite substantial differences between the various domains, progress has been made towards understanding how these estimates are generated and some consistent patterns have emerged. Specifically, prior research has identified two modes of quantitative estimation: *numerical retrieval* and *ordinal conversion* (for a review, see Brown, 2002).

In this paper, we add to the extant literature by identifying a third estimation mode, *feature-based inference* (FBI). Specifically, we argue that people sometimes generate quantitative estimates by combining, in an additive manner, numerical values associated with a small number of critical features. Among the estimation modes identified by this research program, only FBI is a variant of the classic additive compensatory model (Brunswik, 1952, 1956; Hammond, 1955; Hammond, Stewart, Brehmer, & Steinman, 1975). Previous research has demonstrated that models from this family can be applied to a wide variety of judgment and decision contexts (e.g., medical judgment, Wigton, 1996; educational decisions, Heald, 1991; venture capital, Zacharakis & Meyer, 1998; public policy, Adelman, Stewart, & Hammond, 1975; etc.). The present study extends this stream of research and demonstrates that this type of process can also be used when people generate numerical estimates.

2. Estimation modes and seeding effects

People use a variety of strategies to generate numerical estimates. Although strategies can differ from one another in many ways, it has been possible to identify two broad estimation modes which make it possible to classify strategies according to their preconditions and their core processing assumptions (Brown, 2002). One of these is the numerical retrieval mode. A person is said to have used a numerical retrieval strategy when he or she retrieves

* Corresponding author. Tel.: +1 780 248 1091.
E-mail addresses: Kyle.Murray@ualberta.ca (K.B. Murray), Norman.Brown@ualberta.ca (N.R. Brown).
[1] Tel.: +1 780 492 4604.

0001-6918/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.actpsy.2009.05.007

at least one relevant *quantitative fact* while generating an estimate and uses this fact (or these facts) as the basis for a response. In this context, a quantitative fact is defined as explicit pre-existing knowledge of the numerical value posed by a particular item for the current target dimension. "Reconstructive" date estimation (which involves retrieval of landmark dates and temporal inferences based on those dates; Brown, 1990; Friedman, 1993, 2004; Shum, 1998) and rate-based behavioral frequency estimation (which involves the retrieval of rate-of-occurrence facts and computations based on those facts; Conrad, Brown, & Cashman, 1998; Menon, 1993) provide two well-studied examples of numerical retrieval strategies.

The defining element of a numerical retrieval strategy is the recall and utilization of relevant domain-specific quantitative facts. Because these facts are relatively scarce (Nickerson, 1980; Paulos, 1990), the use of these strategies is limited. In contrast, ordinal-conversion strategies appear to be widely applicable. Strategies of this sort typically involve a preparatory stage, called *setting the metric*, during which a response range is defined and partitioned. Once the range has been established, two steps are required to produce a numerical estimate. The first step determines the relative position of the target item or its ordinal value. During the second step, a numerical response is generated by selecting a value from the appropriate portion of the response range (Brown, 2002).

One of the unique features of the research on ordinal conversion is the use of the seeding procedure; Experiments 2 and 3, below, are also seeding experiments. In the typical seeding study, participants first provide numerical estimates for a set of items; then they learn the actual value of at least one of these items and provide a second set of estimates. Seeding experiments have provided evidence that people depend on two independent sources of knowledge when they generate real-world estimates; one of these has been labeled *metric* knowledge (knowledge and beliefs used to define and partition the response range) and the other *mapping* knowledge (knowledge and beliefs used by ordinal processes to determine the relative magnitude of the target items) (Brown & Siegler, 1993; Von Helversen & Rieskamp, 2008).

In addition, these studies provide evidence that seed facts affect estimation performance in different ways. Under most conditions, seeding causes people only to revise their metric beliefs, shifting the range as a whole in the direction indicated and to the degree specified by the metric information carried in the seed value(s). This sort of global updating, called *metric revision*, often results in an across-the-board decrease in absolute error, but no change in the rank-order correlation between estimated and actual values. There are also conditions under which participants redefine the upper and lower values of a portion within the response range – this is called *repartitioning the range*. This occurs when people have an accurate understanding of the response range; when the range is aligned with and portioned to match a well-defined categorical structure; and when people primarily rely on categorical (inheritance-based) inferences to generate their estimates (Friedman & Brown, 2000a, 2000b). This process produces a uniform shift in the post-seeding estimates for items assigned to the revised category (i.e., the category aligned with the revised portion of the response range) and may also produce uniform shifts in items for neighboring categories. (This happens when people assume that neighboring categories are also "strictly adjacent"; Friedman & Brown, 2000b.)

In closing this brief sketch of the estimation literature, it is important to emphasize two points. First, this research indicates that the contents and structure of the relevant knowledge-base play a critical role in determining how people generate their estimates. Second, this work has demonstrated that estimation patterns and seeding effects accurately reflect what people know about a given domain and how they organize this knowledge. The current project took these claims as a starting point and demonstrates that domain knowledge, estimation performance, and seeding effects are explicably related.

3. A feature-based inference model

In this paper we are suggesting that an additive compensatory inference model (e.g., Brunswik, 1952, 1956; Doherty & Kurz, 1996; Hammond, 1955; Hammond et al., 1975) can account for results that other previously identified modes of estimation are unable to adequately explain. This type of model has been used extensively in many different contexts from medical and judicial judgments (e.g., Dhami, 2003; Wigton, 1996) to educational and financial decision making (e.g., Heald, 1991; Zacharakis & Meyer, 1998). Prior research examining numerical estimation has suggested that "people typically use multiple cues to derive an estimate, that they weigh some cues more heavily than others, and that the weight they assign to each cue is a function of its predictive accuracy relative to the predictive accuracy of competing cues" (Brown & Siegler, 1993, p. 530). However, although such an approach clearly has the potential to make a contribution to our understanding of numerical estimation, this line of thinking has not been incorporated into previous models of the estimation process.

Nevertheless, some important domains of numerical information are clearly structured by a few critical features. For example, season and latitude are two critical features that people could use to estimate ambient temperatures. Similarly, in the domain of consumer products, marketers and manufacturers go to great lengths to differentiate the price of their offerings based on a few critical features (e.g., Anderson, de Palma, & Thisse, 1992; Levitt, 1980). Quality is a common example of such a feature. Higher quality is generally believed to be, and often is, correlated with higher prices (e.g., Leavitt, 1954; Rao & Monroe, 1989) – e.g., Rolls Royce versus Kia automobiles. Brand status and country of origin are other examples (e.g., Rolex versus Timex, or Swiss versus Chinese). Other diagnostic features are product-specific such as size for TV screens, processor speed for computers, neighborhood for homes, and so on. The point is that consumer products are often conceived, designed and sold based on the features that they do (or do not) possess. As a result, it is reasonable to suspect that price estimates are a numerical domain in which feature-based inferences play an important role.

In domains structured by key features, the additive value of the features can be critical. That is, although one feature does not provide enough information to allow for a reasonable estimate, two features can be very diagnostic. Consider, for example, estimates of ambient temperature. Knowing the season or latitude alone is not very helpful. However, together they can allow for reasonably accurate estimates (e.g., estimating the temperature in winter at the equator versus at the north pole or in Chicago (42°N) in the summer versus the winter). Similarly, knowing individual features such as the category that a product comes from (e.g., watch, dress shirt, coffee, and SUV) or the brand name (e.g., Armani, Old Navy, Starbucks, and Ford) are useful at only a very general level; however, knowing both pieces of information allows for substantially more accurate price estimates.

The basic premise underlying our model is an expectation that in some domains, such as prices, people can rely on a small number of critical features to generate estimates. Specifically, following this approach, an estimate would be produced by associating a weight with each critical feature and then summing the weighted terms. In many cases, when the to-be-estimated entity possesses none of the critical features its value will still be greater than zero, which implies a constant term in the additive process. Formally, this model can be represented with the following general form:

$$n = \alpha + \sum w_i x_i$$

where n is the numerical estimate; α is the constant value assigned to all entities that possess none of the i critical features; w_i represents the weights assigned to each feature i, indicating the amount by which n increases for each increment of one in the value of the feature x_i.

As an example of the FBI process, consider the domain of new automobiles and an individual who believes that a new vehicle costs at least $15,000 (i.e., α = $15,000). Also, assume s/he uses two critical features, brand image and type – that is, vehicles have either a premium brand image (value = 1) or not (value = 0) and are of a premium type (e.g., SUV or sports car; value = 1) or not (e.g., basic passenger car; value = 0). These features are then weighted – for example, w_{brand} = $10,000 and w_{type} = $20,000. With this information an estimate can be produced for any vehicle by simply deciding whether or not it is a premium brand and/or a premium type of automobile. A vehicle that is believed to be premium on both features would be estimated to cost $45,000 (i.e., $n = \$15,000 + (\$10,000 \times 1) + (\$20,000 \times 1)$); a vehicle that is not perceived as premium on either feature would be assigned a value of $15,000 (i.e., $n = \$15,000 + (\$10,000 \times 0) + (\$20,000 \times 0)$); the price of a vehicle that is from a premium brand, but is a basic type of automobile, would be estimated at $25,000 (i.e., $n = \$15,000 + (\$10,000 \times 1) + (\$20,000 \times 0)$); and a vehicle that is not a premium brand, but is a premium type of vehicle, would be estimated at $35,000 (i.e., $n = \$15,000 + (\$10,000 \times 0) + (\$20,000 \times 1)$).

It is worth noting that this process is distinctly different from both numerical retrieval and ordinal conversion. As compared to numerical retrieval, people using FBI are not retrieving the price of the target from any other "reference" or "landmark" vehicle. Instead the car is "decomposed" into critical features; the feature values are accessed, weighted and combined to generate a response. FBI is also a very different process than ordinal conversion. Using FBI there is no need to make an ordinal judgment or map an ordinal value onto the response range. In the FBI model, the metric information is determined by the weights associated with each of the critical features.

However, as defined above, the FBI process also includes an estimate of the minimum value (i.e., α) for the entity in question. This value is not directly related to any feature or combination of features, but instead represents the lower bound of the range of estimates. As in the previous vehicle example, α sets a lower bound for estimates (e.g., a price of $15,000) – that is, a vehicle that is not perceived to be "premium" on either critical feature would be assigned a value of $15,000 (i.e., $n = \$15,000 + (\$10,000 \times 0) + (\$20,000 \times 0)$).

3.1. The split-seed hypothesis

A novel prediction of the FBI model, relative to numerical retrieval and ordinal conversion, is that metric information can be stored and revised at the level of critical features. As a result, when people are using an FBI process, they should respond to a seed fact by splitting the information carried by the seed between the critical features. Therefore, the degree of post-seeding revision should be directly related to the number of these features that the seed and the transfer items have in common – i.e., the split-seed effect. Specifically, when a seed fact (e.g., automobile price) that is incongruent with the first set of estimates is made available, the second set of estimates should be revised in a manner that is dependent upon the relationship between the seed fact, the critical features and the automobile being estimated. Therefore, continuing with the above example, vehicles that share both of the critical features will be affected the most, followed by those that share a single critical feature, with the smallest effects being reserved for the automobiles that do not share those features with the seed vehicle.

For example, after learning that one automobile (that is both a premium brand and a premium type) is $15,000 more expensive than it was originally thought to be, an individual might adjust the weights assigned to the critical features such that premium brand automobiles are now estimated to be $5000 more expensive and premium type automobiles are estimated to be $10,000 more expensive. "Splitting the seed" in this manner suggests that all premium brand automobile estimates will increase by $5000; all premium type automobile estimates will increase by $10,000; and estimates for automobiles that share both critical features (i.e., premium brand and type) will increase by $15,000. Neither numerical retrieval nor ordinal conversion would predict or easily account for these effects. The split-seed hypothesis is tested directly in Experiments 2 and 3. First, however, it is necessary to identify the features that are critical in structuring automobile price knowledge.

4. Experiment 1

The features used to generate estimates will, obviously, vary between domains (Brown & Siegler, 1993); therefore, which features play what role within a particular domain of interest is an empirical question (Brown, 2002). For this reason, Experiment 1 is a pilot study that examines the relative efficacy of different features in producing automobile price estimates. That information is then used to design Experiments 2 and 3, and to produce more detailed predictions that allow us to further test the value of this model.

4.1. Method

4.1.1. Participants

Data were collected from 22 undergraduate psychology students (11 male and 11 female) at a large Canadian university (each student was paid $5 for participating). The mean age was 22.3 years and ranged from 19 to 29. Of these participants 68% currently owned a vehicle, 41% were planning to buy a vehicle in the next 12 months, and 18% had purchased a vehicle in the past 12 months. These participants reported that they have thought about vehicles an average of 11.2 times over the past month, with a range of 1–45 times.

4.1.2. Design

All participants were asked to provide price estimates for 18 automobiles. The vehicles used in all three experiments reported in this paper were new models when the data were collected. Dependent variables were estimates of the *base price* – i.e., the minimum price at which the vehicles can be acquired – for each of the 18 automobiles and ratings of perceived quality for each of the six brands.

4.1.3. Materials

The 18 automobiles for which participants estimated prices in this experiment were taken from six brands (Ford, Chevrolet, Toyota, Honda, BMW, and Mercedes) and three product categories (passenger cars, sports cars, SUVs). The models used were Passenger cars: Ford Focus, Chevrolet Cavalier, Toyota Echo, Honda Civic, BMW 323i, Mercedes C230; Sport Utility Vehicles (SUVs): Ford Explorer, Chevrolet Blazer, Toyota 4Runner, Honda CRV, BMW X5, Mercedes M-Class; Sports Cars: Ford Mustang, Chevrolet Corvette, Toyota Celica, Honda S2000, BMW Z3, Mercedes SLK-Class. The participants were provided with the brand and the model (e.g., Honda S2000 or BMW 323i) for each vehicle for which they were asked to provide a price estimate.

4.1.4. Procedure

All respondents performed a familiarity rating task, a price estimation task, and a brand quality rating task. All the tasks were completed on a personal computer in a laboratory with the experimenter present. Participants were first asked to complete a familiarity ratings task. Each participant was asked to rate their level of familiarity with each of the 18 automobiles. The automobiles were each presented only once and on the computer screen one at a time; each participant provided a rating, on a 0 (not at all familiar) to 9 (very familiar) scale, before continuing onto the next vehicle. Participants were given the brand and model of the vehicle followed in parentheses by the category – for example, Ford Explorer (SUV). The vehicles were presented in random order; however, each of the three product classes (SUVs, sports cars and passenger cars) was presented before any of the three classes were repeated.

After completing the familiarity task, participants provided base price estimates for each of the 18 automobiles. The automobiles were again presented in the same randomized fashion as in the familiarity rating task. After providing price estimates participants were asked to provide quality ratings, on a 0 (very low quality) to 9 (very high quality) scale, for each of the automobile brands, with the brands being presented in a random order. At the end of the experiment all participants completed a survey that asked whether or not they owned a vehicle, whether or not they planned to buy a vehicle in the next 12 months, whether or not they had purchased a vehicle in the previous 12 months, as well as the participant's gender. We also asked participants: How many times in the past month have you thought about or discussed vehicles?

4.2. Results

Two measures were computed for each participant and each estimate: signed normalized error (SNE = [estimated price − actual price]/actual price) and absolute normalized error (ANE = |SNE|). These measures were averaged across the six instances representing each vehicle class yielding three SNEs and three ANEs. Three estimate means, one for each vehicle class, were also computed. In addition, we computed separate rank-order correlations between estimated and actual price for each vehicle class, along with the percentage of correct estimates – consistent with previous studies (see Monroe & Lee, 1999) a correct response was considered to be ±5% of the current price of the to-be-estimated vehicle. Finally, a single rank-order correlation was computed over all 18 responses for each participant.

Percentage correct can be seen as providing an upper bound on the use of a price retrieval strategy. The rank-order correlations indicate whether participants had a good sense of the relative prices of the vehicles regardless of the accuracy of their metric knowledge (Brown & Siegler, 1993). ANE provides a normalized measure of overall accuracy; ANE is small when estimates tend to be accurate and large when they do not. SNE provides a normalized measure of bias; a negative SNE indicates a bias to underestimate prices and a positive SNE a tendency to overestimate them. To facilitate comparisons to previous research, which focused on accuracy within product categories, in this section we restricted our analysis to product class. This approach reflects the exploratory nature of this experiment. As mentioned above, one of the primary goals of this experiment was to obtain evidence that would allow us to identify key features that structure people's knowledge of automobile prices.

Table 1 reports the hit-rate or percentage of correct estimates. Consistent with previous work on consumer price knowledge our data indicate that exact price recall is rare. Across the three product categories accuracy ranged from 10.6% (for Sports Cars) to 12.1% (for SUVs and Passenger Cars). To test for differences in price estimation accuracy between the three product categories we used a random effects logistic regression model, with accuracy coded as: 1 = accurate price estimate (within ±5% of the true price); and 0 = inaccurate price estimate. We find no significant difference in accuracy between the product categories ($z = 1.07$; $p = 0.283$). This

Table 1
Accuracy of mean estimates (SNE and ANE calculations) and Spearman rank order correlation coefficients (for the first set of price estimates and true base prices) across all three experiments.

	Sports cars			SUVs			Passenger cars			All vehicles		
	Experiment 1	Experiment 2	Experiment 3	Experiment 1	Experiment 2	Experiment 3	Experiment 1	Experiment 2	Experiment 3	Experiment 1	Experiment 2	Experiment 3
Percent accurate – first estimates	10.6%	8.4%	10.2%	12.1%	13.7%	12.6%	12.1%	11.6%	11.6%	11.6%	11.2%	10.3%
SNE – first estimates	−0.015	0.022	**−0.179**	−0.054	0.003	−0.025	**0.156**	**0.273**	0.059	0.029	**0.102**	−0.023
SNE – second estimates		**0.258**	0.046		**0.231**	**0.394**		**0.541**	**0.290**		**0.349**	**0.309**
ANE – first estimates	**0.336**	**0.348**	**0.576**	**0.257**	**0.267**	**0.403**	**0.261**	**0.406**	**0.465**	**0.285**	**0.332**	**0.448**
ANE – second estimates		**0.419**	**0.465**		**0.330**	**0.341**		**0.598**	**0.371**		**0.451**	**0.369**
Rank order correlations	0.485	0.468	0.267	0.703	0.691	0.279	0.734	0.801	0.646	0.641	0.653	0.443

Note: **Bold** text indicates a SNE or ANE value that is significantly different from 0; with p-values < 0.01 (based on a one-sample t-test). All other SNE and ANE values are not significantly different from 0; with p-values > 0.10.

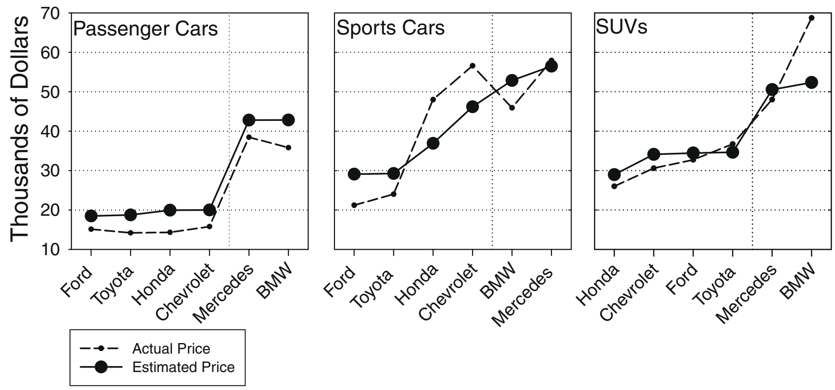

Fig. 1. *Experiment 1* – mean price estimates and true prices across the three product categories.

level of recall is lower than other recent studies of exact price recall accuracy, measured as ±5% of the true price, which ranged from 24% to 65% (see Monroe & Lee, 1999, for a review). In addition to the relatively low levels of accuracy, the absolute magnitude of the estimation errors is nontrivial. Table 1 reports the mean absolute normalized error (ANE) calculations for each product class. When expressed as a percentage, the ANEs range from 25.7% (SUVs) to 33.6% (Sports Cars), which is higher than a range of 6.0–19.5% previously reported for a number of different product categories (Monroe & Lee, 1999). To test for differences in ANEs between product categories we used a random effects GLS regression model (we use a random effects GLS regression model to account for the fact that each participant provides a price estimate for each of the 18 automobiles). We find a significant difference ($z = -2.490$, $p = 0.013$), which indicates that the ANEs are significantly larger for sports cars than for the other vehicle categories.

Prior research in this area has suggested that scalar variability can be important in understanding estimate accuracy, especially in the domain of prices (Grewal & Marmorstein, 1994; Marques & Dehaene, 2004). In particular, the Weber fraction – defined as the standard deviation of estimates across participants divided by the mean of the estimates – has been shown to be a useful measure of price estimation accuracy (Dehaene & Marques, 2002). Across all automobile estimates in this study, the Weber fraction is 0.49. This is higher than the 0.34 value reported by Dehaene and Marques (2002) across a broader range of products, which is consistent with the above analysis suggesting that automobile price estimates are less accurate than for other types of products.[2]

Although price recall is less accurate and the magnitude of the error is larger for automobiles than has been reported for other product categories, this finding is consistent with the broader literature on numerical estimation that reports hit rates (estimates that are ±5% of the actual value) that range from 0% to 23% and absolute normalized estimation errors (ANEs) that range from extremely accurate (<0.01) to estimates that are incorrect by orders of magnitude (Brown, 2002). In addition, because the prices of commonly purchased products are generally more accurately recalled than the prices of infrequently purchased goods and recreational products (Estelami, 1998), it is reasonable that estimation errors for automobile prices are higher than for other products from more commonly purchased categories, such as groceries. Our SNE calculations (Table 1) indicate that while there is a tendency to over-estimate the price of the Passenger Cars, there is no general tendency towards over- or under-estimation in the Sports Car or SUV categories. Based on a random effects GLS regression model, there is a significant difference in the mean SNEs between the product categories ($z = 4.310$; $p < 0.001$), which confirms the apparent bias towards over-estimation in the passenger car category.

In brief, participants were rarely accurate in terms of the percentage of correct estimates, and in absolute terms they misestimated the price of the automobiles by a substantial margin. However, as predicted, they were relatively accurate in terms of their rank ordering of the automobiles by price (Table 1 reports the mean individual-level average rank-order correlations by product class). For all vehicles the rank order correlation coefficient is 0.641, which is consistent with the range for other product categories reported by Conover (1986) of 0.429–0.773. The three price profiles (i.e., mean estimates by product class) are presented in Fig. 1. Across all three of the profiles it is evident that the products are stratified into two levels of brand status – a basic tier and a luxury tier. The basic tier contains Fords, Chevrolets, Toyotas and Hondas, while the luxury tier is composed of the BMWs and the Mercedes.

To more formally examine the grouping of the price estimates, a series of hierarchical cluster analyses using Ward's method and the squared Euclidean distance measure was performed on the individual-level price estimates within each product class. Table 2 illustrates the optimal cluster analysis solutions for each of the categories, and provides additional evidence that the estimates are grouped into two levels of brand status (rather than three or more) within each product class. There is one obvious exception to the normal brand status grouping: the Chevrolet Corvette (see Table 2), which is an outlier with regard to the basic price tier (see Fig. 1, Sports Cars). However, when participants are divided into two groups based on a median split of the Corvette familiarity ratings the mean Corvette price estimates for the two groups differ (one-tailed *t*-test: $t_{(21)} = -1.6212$, $p = 0.07$). Participants in the high familiarity group estimate the Corvette to be $45,625 (*SD* = $23,463), while participants in the low familiarity group estimate the Corvette to be $29,500 (*SD* = $11,362). Interestingly, while the mean for the familiar group is closer to the true price for the Corvette ($56,585), the mean for the unfamiliar group is very consistent with the average price estimate for the basic sports car tier (*M* = $30,833, *SD* = $12,937). A median split of the price estimates based on the familiarity ratings does not result in any significant differences between the two groups for any of the other 17 vehicles (all $p > 0.10$).

Based on the cluster analysis and the pattern of mean price estimates evident in Fig. 1, it appears that product class (SUV, Sports Car, Passenger Car) and brand status (luxury versus non-luxury) are the features that the price estimates are based on. We tested

[2] Other important findings from this line of research are not directly applicable here. For example, we do not find that the Weber fraction increases as price increases, which likely reflects the fact that all of our stimuli are automobiles with prices that are all much higher than other types of consumer products. Similarly, we do not find any correlation between the Weber fraction for estimates of specific automobiles and familiarity ratings for those same vehicles ($r = 0.023$).

this possibility by comparing the ability of these two features to predict participant's price estimates with an alternative model that includes other potentially important predictors: perceived quality, familiarity with the vehicle, brand name, country of origin, gender, whether or not the participant owns a vehicle, whether or not the participant plans to buy a vehicle in the next 12 months, whether or not the participant bought a vehicle in the previous 12 months, and the number of times participants had thought about automobiles in the past month. One advantage of the FBI model that we are proposing is that it lends itself directly to regression analysis, which has been well established as an appropriate method for testing feature-based judgment theories similar to the one that we are advocating here (Doherty & Brehmer, 1997; Hammond, 1986). In this case, the potentially important features of the environment serve as independent variables, while the numerical estimates themselves are the dependent variables. Specifically, we use a random effects GLS regression model, consistent with other linear compensatory models of judgment (e.g., Hammond et al., 1975; Payne, Bettman, & Johnson, 1992, 1993), with the following general form:

$$Y_e = \beta_0 + \beta_i x_i + v_i + \varepsilon_{ij}$$

where Y_e is the participants' price estimates, β_0 is the intercept, β_i is the weight given to the critical feature i (either CLASS or STATUS). We dummy coded CLASS as 1 (premium category: SUVs and Sports Cars) or 0 (non-premium category: Passenger Cars), and STATUS was dummy coded as 1 (luxury tier) or 0 (non-luxury tier). Log price estimates are used rather than absolute price estimates, because the distribution of estimates for this data set is truncated (with a minimum estimate of \$10,000 and a maximum estimate of \$100,000). To account for the fact that each participant makes 18 price estimates, two error terms are included in the model. The subject-specific residual (v_i) differs between subjects, but its value is constant for any particular subject. The general error term is represented by ε_{ij}.

As expected the two-feature (product class and brand status) model of participants' price estimates explains the majority of the variance in the data ($R^2 = 0.582$). Adding in all the remaining predictors has only a small effect on overall fit of the model (full model $R^2 = 0.599$). The intercept ($z = 185.84$, $p < 0.001$, $SE = 0.053$) is significant, as are the coefficients for product class ($z = 15.02$, $p < 0.001$, $SE = 0.029$) and brand status ($z = 17.10$, $p < 0.001$, $SE = 0.029$). Therefore, the two-feature additive inference model of participants' price estimates (with the intercept and the coefficients expressed in dollars) is

$$Y_e = \$18,868 + \$10,400\ (\text{CLASS}) + \$12,228\ (\text{STATUS}).$$

4.3. Discussion

A starting point for this experiment was to determine how accurate people are when they estimate automobile prices. On average, participants in Experiment 1 were relatively inaccurate as compared to those of earlier studies of consumer price knowledge (Monroe & Lee, 1999). This low hit rate is consistent with the assumption that people do not directly recall specific product prices. It might also reflect two important facts about automobiles: (1) market prices tend to fluctuate on a regular basis (between regions and between dealerships within a region); and (2) consumers tend to buy a new vehicle relatively infrequently (and so update their knowledge structures less frequently than they might for other products). Nevertheless, although we observe price knowledge that is much less accurate than would be predicted by normative economic models, our participants exhibit levels of recall accuracy that are high relative to other domains of quantitative estimation (Brown, 2002). This may reflect the more frequent exposure people have on prices as compared to other quantitative facts such as national populations or the nutritional value of fast foods.

One reason we chose automobiles as our stimuli was that information about vehicle prices, especially for the familiar brands that we have chosen, is relatively common. Therefore, it is not surprising that the vehicles that were most poorly estimated (e.g., Honda S2000 and BMW X5) are more niche products with relatively small market shares and prices that are atypical within their product class. The true prices for these vehicles deviate dramatically from the other brands with which they are normally grouped. Nevertheless, the absolute magnitude of the observed estimation errors is large enough to deserve attention, as previous research has demonstrated that such errors can have a substantial impact on product demand and consumers' perceptions of transaction value (Putler, 1992; Thaler, 1985).

It should be noted that we are not suggesting that consumers are never able to recall specific prices. Clearly, there are some familiar products for which consumers are able to recall an accurate price (Estelami, 1998). However, we argue that as in other domains of numerical estimation, direct recall is the exception rather than the rule (Brown, 2002). In most cases, consumers are forced to rely on more general knowledge to construct price estimates. The consistent stratification of price estimates into two distinct price tiers is evidence of an estimation strategy based on a general level of price knowledge – i.e., a product class and brand status inference process.

Table 2
Experiment 1: optimal cluster analysis solutions.

	Cluster number		
	Sports cars	Sport utility vehicles	Passenger cars
Ford	1	1	1
Chevrolet	3	1	1
Honda	1	1	1
Toyota	1	1	1
BMW	2	2	2
Mercedes	2	2	2

5. Experiment 2

Experiments 2 and 3 compare our feature-based inference model with other potentially relevant theories by examining the predictions they make about how estimates will be updated when respondents are presented with relevant new information.

5.1. The split-seed hypothesis for automobile price estimates

We predicted that when people are presented with a seed price that is incongruent with their first set of estimates, they will update their second set of estimates by adjusting the weight that is given to the critical features of brand status and product class. For example, when we presented respondents with the true price of the BMW X5, which is much more expensive than other SUVs, we expected that would have a stronger effect on other luxury SUVs than it would have on SUVs in general, other luxury vehicles, or other non-luxury non-SUV automobiles. That is, the effect of the seed will be split depending on the relevance of the change in the weight given to the critical features with respect to the automobiles for which an estimate is being made.

5.2. Alternative predictions

However, it is not necessarily the case that people will react this way to a seed price. In fact, other well-known theoretical frameworks predict very different effects. To begin our examination of these alternative predictions, we look at the neo-classical theory of economics (Marshall, 1890; Monroe & Lee, 1999), which predicts that prices are known for all goods and that knowledge of one product's price is independent of other prices for related products. In this case, when asked to estimate automobile prices, people would employ a numerical-retrieval strategy and simply recall the prices of each vehicle from memory. As a result, the only estimate that would be revised in the second set would be the price of the seed vehicle. Although this may seem to be a relatively weak alternative given the results that we have discussed from prior studies of estimation, the neo-classical theory continues to be an important foundation in many current studies of economic behavior (e.g., Becker, 1993; Lucas, 1987; McFadden, 1999), including studies of price knowledge (e.g., Briesch, Krishnamuthi, Mazumdar, & Raj, 1997; Kalyanaram & Little, 1994; Putler, 1992).

Another possibility is that all the automobiles will shift by a constant amount as the overall metric is reset, while the distance between the ordered entities is kept constant. This type of shift is consistent with revising the metric (i.e., the range of values) within the ordinal conversion process. For example, in studies of subjective geography, it has been found that information that changes a participant's estimate of one city's latitude can affect all other cities by a constant amount (Friedman & Brown, 2000a, 2000b) – e.g., when a participant realizes that s/he was 10° too low in his or her estimate of the location of New York, the entire distribution is uniformly revised upwards by 10°. An automobile seed fact could have the same effect. After learning that his or her estimate for the BMW X5 was $20,000 too low, the respondent could increase all estimates in the second set by $20,000. The metric is revised, but the ordering of and distance between entities within the distribution remains constant.

Finally, it may be that, consistent with the ordinal-conversion process, people respond to a seed price by repartitioning the range. One way to repartition the range is to move sub-categories within the range, independent of one another. For example, in studies of subjective geography, people group city latitudes in some regions together and when a seed fact is presented it affects the category it is from, while cities in other categories are not affected (e.g., Friedman & Brown, 2000a, 2000b). This type of categorical knowledge structure implies that, for example, after exposure to the BMW X5's true price, participants' would revise all estimates for other luxury vehicles and SUVs (the sub-categories that the BMW appears to belong to based on the results of Experiment 1) to keep prices within the categories consistent. However, there would be no reason to revise price estimates that were not part of the SUV or luxury categories. More importantly, all the categories that are affected would move by a constant amount. Therefore, continuing with the BMW X5 example, the adjustment in the SUV category would be the same as the adjustment in the luxury vehicle category and it would also be the same for vehicles that were in both the luxury and SUV categories – i.e., there would be no additive effect for vehicles that are members of both categories.

5.3. Method

5.3.1. Participants

Data were collected from 53 undergraduate students (17 male and 36 female) in psychology and business at a large Canadian university. The mean age was 19.9 years and ranged from 17 to 38. Of these participants 42% currently owned a vehicle, 21% were planning to buy a vehicle in the next 12 months, and 17% had purchased a vehicle in the past 12 months. These participants reported that they have thought about vehicles an average of 14.4 times over the past month, with a range of 0–100 times.

5.3.2. Design

All participants were randomly assigned to one of three seed conditions. Twenty participants in Condition 1 were presented with the price for the Honda S2000 sports car ($48,000). Seventeen participants in Condition 2 were presented with the seed price for the BMW X5 sport utility vehicle ($68,700). Sixteen participants in Condition 3 were presented with the seed price for the Lexus GS400 sports car ($67,960). Participants were asked to provide price estimates for 21 automobiles. The dependent variable was the price estimate for each of the 21 automobiles.

5.3.3. Materials

The 21 automobiles for which participants estimated base prices in this experiment included the same 18 vehicles from Experiment 1 plus an additional three vehicles from the Lexus brand (Lexus ES300 passenger car, Lexus GS400 sports car, and the Lexus RX300 SUV).

5.3.4. Procedure

The first part of this experiment was a replication of Experiment 1, with the same three tasks and with the order of stimuli presented on the computer screen one at a time and in random order (as in Experiment 1). However, as discussed above, a new brand (Lexus) was added to the stimuli, thereby bringing the total number of vehicles to 21. The Honda S2000 and BMW X5 seeds were chosen because participants were very inaccurate in their price estimates for these cars in Experiments 1. The Lexus GS400 seed fact was chosen because its true price is atypical, which we have argued reduces accuracy in price estimates. Including this vehicle in the set allowed us to test the prediction that its price would be substantially under-estimated. After the seed fact was presented, participants were asked to again estimate the price of the 21 automobiles. This time, the seed fact (the model name and the model price for one of the Honda, Lexus or BMW) was present in the right hand corner of the screen for participants to see as they made their estimates. At the end of the experiment all participants were asked to complete the same survey as in Experiment 1.

5.4. Results

The initial price estimates from Experiment 2 are plotted in Fig. 2. This figure illustrates that the results of Experiment 2 replicate the results of Experiment 1, with respect to the pattern and grouping of initial price estimates. Price recall accuracy, the magnitude of the estimation error and the accuracy of participants' rank order correlations are reported in Table 1. The general patterns of price estimation found in Experiment 1 are replicated in Experiment 2. In addition, as predicted, the price of the Lexus GS400 was substantially underestimated (one sample *t*-test: $t_{(52)} = -11.483$, $p < 0.001$). The true price of the Lexus is $67,960; therefore, the mean pre-seed estimate was incorrect by $22,526 or 33%.

We find that, on average, the second set of price estimates was revised in all three conditions, as participants reconcile their initial underestimate for the seeded vehicle with the seed price they are given (see Fig. 3). To test the statistical significance of the revisions between the first and the second set of price estimates we use a random effects GLS regression model with the log of the individual-level price estimates as the dependent variable and an independent variable dummy coded as 0 = first estimate and 1 = second estimate. We find that the difference between the first and second estimates is significant ($\beta = 0.223$, $z = 11.95$, $p < 0.001$). Converting the β coefficient into dollars indicates that the second set of estimates increased by an average of $8224. The first and second price estimate means, as well as the standard deviations, are reported in Table 3 by experimental condition and product class.

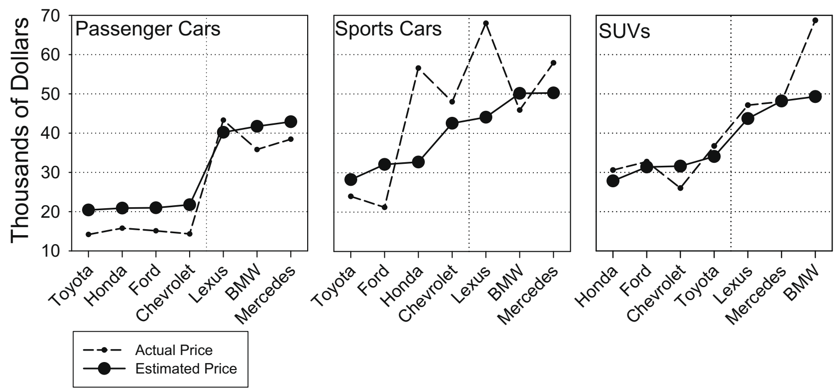

Fig. 2. *Experiment 2* – mean initial (pre-seed) price estimates and true prices across the three product categories.

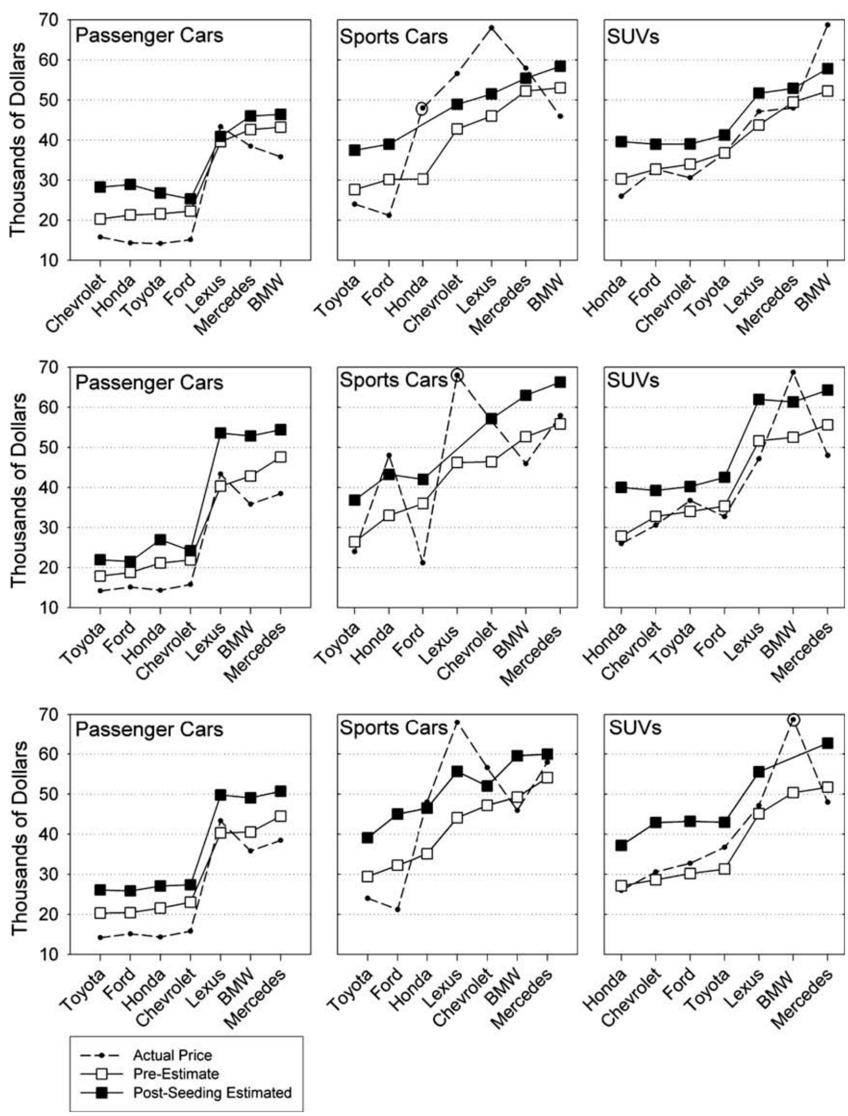

Fig. 3. *Experiment 2* – mean first and second price estimates, and true prices, across the three product categories and the three seed conditions (the seed price is indicated by a circled data point).

Next, we look at the ability of the GLS regression model used in Experiment 1 to predict price estimates on the basis of two-key features: product class and brand status. We apply the same model to the data in each of the seed conditions, as well as to the full set of data collapsed across seed conditions. The results (reported in Table 4) indicate that the two-feature model explains much of the variance in both sets of price estimates, in each condition and collapsed across conditions, in the second experiment. As in the first experiment, adding in the other potentially important predictors does not have a substantial impact on the fit of the model.

Table 3
Experiment 2: mean estimates and standard deviations by condition and vehicle type, first and second estimates (the first and second estimates for the seed vehicle have been removed from the data in each condition).

		First set of estimates				Second set of estimates			
		Basic		Luxury		Basic		Luxury	
		Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Condition 1 (Honda S2000 $48,000)	Passenger cars	$21,363	$7858	$41,800	$15,445	$27,313	$9417	$44,433	$13,062
	Sports cars	$33,517	$15,437	$50,400	$18,068	$41,767	$13,634	$55,100	$11,040
	SUVs	$33,438	$12,929	$48,467	$17,929	$39,700	$10,180	$54,133	$15,488
Condition 2 (BMW X5 $68,700)	Passenger cars	$21,309	$6701	$41,765	$17,865	$26,574	$10,401	$49,843	$16,435
	Sports cars	$35,985	$15,355	$49,118	$17,219	$45,647	$16,778	$58,373	$16,404
	SUVs	$29,294	$7521	$48,412	$17,978	$41,544	$12,049	$59,118	$12,426
Condition 3 (Lexus GS400 $67,960)	Passenger cars	$19,906	$9004	$43,563	$14,233	$23,641	$11,367	$53,583	$10,518
	Sports cars	$35,453	$14,622	$54,219	$13,090	$44,797	$14,344	$64,594	$12,000
	SUVs	$32,469	$10,753	$53,250	$14,246	$40,484	$11,317	$62,500	$10,518

Table 4
Experiment 2: the two-key features model applied to each condition and collapsed across conditions.*

	Estimate set	Constant	Class	Status	R^2 (two-key features)	R^2 full (all 11 predictors)
Full model (collapsed across conditions)	First	$21,164	$8580	$14,363	0.484	0.527
	Second	$26,388	$11,973	$15,087	0.511	0.547
Condition 1 (Honda S2000 – $48,000)	First	$21,789	$7944	$13,619	0.487	0.554
	Second	$27,203	$10,398	$12,175	0.428	0.488
Condition 2 (BMW X5 – $68,700)	First	$21,571	$7690	$13,513	0.400	0.443
	Second	$27,198	$12,260	$14,794	0.512	0.545
Condition 3 (Lexus GS400 – $67,960)	First	$19,964	$10,363	$16,278	0.580	0.608
	Second	$24,500	$13,836	$19,412	0.628	0.632

*All the intercepts, as well as all the coefficients for product class and brand status, reported in the table are significant for both the first and the second set of estimates (all $p < 0.001$).

Follow-up tests provide additional evidence that the seed price has an effect that differs across vehicles in a manner that is consistent with the split-seed hypothesis, but inconsistent with the alternative explanations. For example, in Condition 1 (the Honda S2000 sports car seed price), luxury passenger car price estimates are revised by an average of $2633, which is significantly less than the average revision to basic passenger cars (M = $5950; $F_{(1,138)} = 4.243$, $p = 0.041$). In Condition 2 (the BMW X5 SUV seed price), luxury sports cars are revised by an average of $9662, which is significantly more than the average revision to basic passenger cars (M = $5265; $F_{(1,134)} = 3.912$, $p = 0.037$). The results in Condition 3 are similar, as luxury SUVs (M = $9250) are revised by an amount equivalent to luxury sports cars (M = $10,375; $F_{(1,78)} = 0.130$, $p = 0.720$), but significantly greater than basic passenger cars (M = $3735; $F_{(1,110)} = 3.927$, $p = 0.005$).

Consistent with the results of other studies on the effect of seed information on second estimates, the seed prices do not act as traditional anchors (Brown & Siegler, 2001). That is, in addition to second price estimates assimilating towards the seed value (e.g., lower estimates increasing towards seed), some price estimates are moving away from (contrasting with) the seed price. Fig. 3 illustrates that when the price of the Honda S2000 is the seed fact the second set of price estimates contrast (i.e., estimates increase away from the seed) for luxury sports cars (first estimate mean = $50,400; second luxury tier estimate mean = $55,100; $t = -2.955$, $p = 0.002$) and for luxury SUVs (first estimate mean = $48,466; second estimate mean = $54,133; $t = -4.357$, $p < 0.001$).

5.5. Discussion

Overall, the results from Experiment 2 provide additional evidence in favor of a general knowledge structure for automobile prices, organized by brand status and product class. When presented with an external reference price, that is inconsistent with their initial price estimates, participants revise the range and the mean of their estimates; yet the general ordinal structure is maintained. While Experiment 1 indicated that product class and brand status are critical features for generating price estimates, the two-feature inference model could only be fit ex-post. Experiment 2 replicates the results of Experiment 1, and finds that the proposed FBI model fits the data well for both the pre- and post-seed price estimates.

In addition, Experiment 2 tested the split-seed prediction of the FBI model against a set of alternative theories. The results clearly indicate that a seed fact specific to one automobile affects price estimates for other automobiles. This finding is inconsistent with the neo-classical theory, which contends that price knowledge is product-specific and that revising price knowledge for one product should not affect other products. The results of the second experiment also indicate that presenting participants with a seed price does not affect all other price estimates equally, which rules out the possibility that price knowledge is revised by a constant amount.

Although the revised estimates clearly result in a change to the range and mean of the price estimates, this change is not consistent with a simple repartitioning of the range based on category assignment because all vehicles within a particular category (e.g., luxury or SUV) are *not affected by the same amount*. Instead, automobiles that share both features with the seed vehicle are affected more than those that share only a single feature or no features. That is, brand status and product class have significant independent effects on price estimates.

The seed price also has an effect on the lower bound of the range (Table 4). This suggests that the seed caused people to revise their metric knowledge of α, in addition to revising the weights assigned to each critical feature. This finding is consistent with previous work on numeric estimation, which has found that a revision of global metric knowledge is a common effect of seeding the knowledge base (Brown, 2002; Brown & Siegler, 1993). As in ordinal conversion, which begins with a setting of the metric followed by a mapping of values into that revised range, feature-based integration also involves a general resetting of the metric. However, the FBI model is distinct from ordinal conversion in the *process* by which individual entities are assigned values. While FBI predicts and explains the split-seed effect – that automobiles that share both features with the seed vehicle are affected more than those that share only a single feature or no features – ordinal conversion cannot account for these results.

6. Experiment 3

In this third experiment we elicit estimates in different contexts, that do not include all product classes and levels of brand status, which allows us to address the possibility that the estimates we have seen so far are being driven by the context that we have given to participants. In other words, it is possible that the stratification we have seen and the means for the different clusters of estimates may be the result of the specific estimation context rather than the participants' true knowledge structure for automobile prices. For example, passenger cars are seen as the least expensive vehicles and so are priced at the lower bound of the range (approximately \$18,000). It is conceivable that if we asked only about luxury and basic SUVs, participants would produce basic SUV estimates at the lower bound of the range (approximately \$18,000) because they are the least expensive vehicles in the given context. This possibility is examined in Experiment 3.

6.1. Method

6.1.1. Participants

Data were collected from a total of 63 undergraduate students (29 male and 34 female) majoring in business at a large Canadian university, who were participating to earn course credit. The mean age was 21.7 years and ranged from 19 to 32. Of these participants 56% currently owned a vehicle, 27% were planning to buy a vehicle in the next 12 months, and 6% had purchased a vehicle in the past 12 months. These participants reported that they have thought about vehicles an average of 5.2 times over the past month, with a range of 0–50 times.

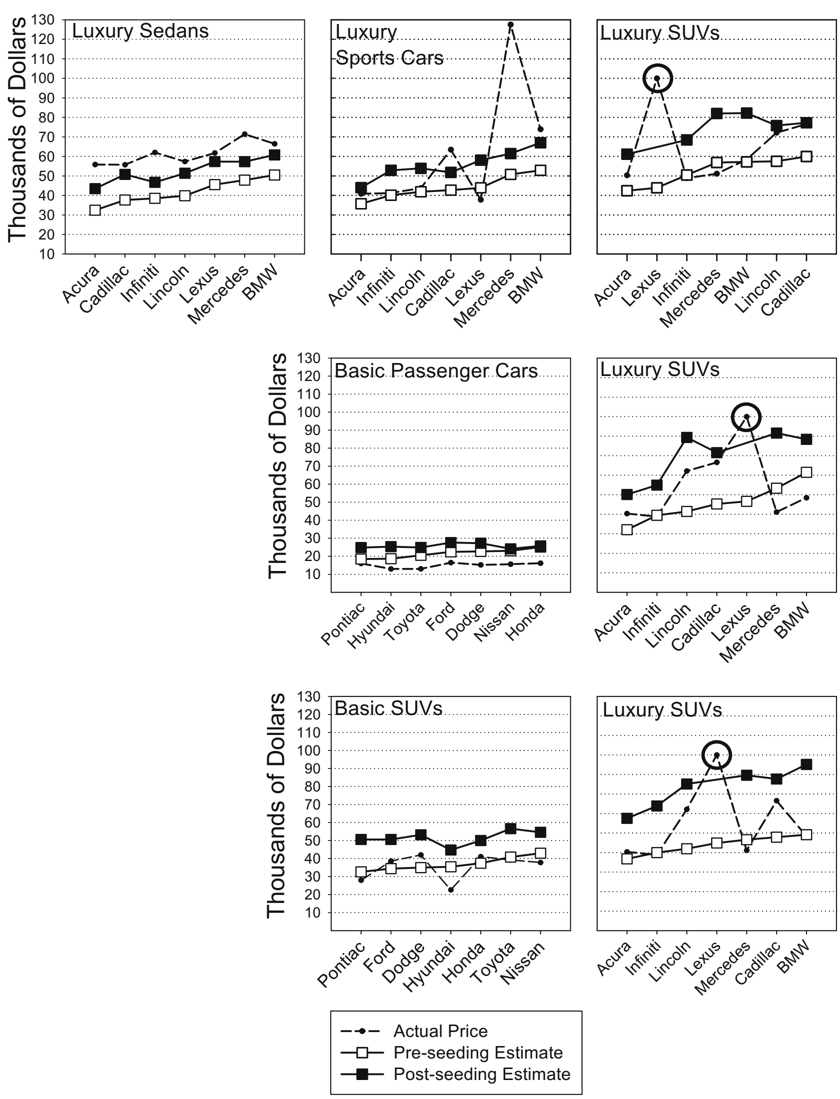

Fig. 4. *Experiment 3* – mean first and second price estimates, and true prices, across the three experimental conditions (the seed price is indicated by a circled data point).

6.1.2. Design

All participants were randomly assigned to one of three experimental conditions. Each condition varied the make-up of the set vehicles to be estimated (see Appendix A). In Condition 1, participants were asked to estimate the base price for 21 vehicles from the luxury SUV, luxury sports sedan and luxury passenger car categories. In Condition 2, participants were asked to estimate the base price of 14 automobiles selected from the luxury SUV and basic passenger car categories. In Condition 3, participants were asked to estimate the base price of 14 automobiles from the luxury and basic SUV categories.

6.1.3. Materials

The automobiles for which participants estimated prices in this experiment are listed by experimental condition in Appendix A.

6.1.4. Procedure

Following the procedure in Experiment 2, each participant was asked to estimate the base price for a series of automobiles twice. Vehicles were presented using the same randomization procedure employed in Experiments 1 and 2. The second time participants generate price estimates they do so after having been told the price for the Lexus LX470 ($99,950.00). At the end of the experiment all participants were asked to complete the same survey as in Experiments 1 and 2.

6.2. Results

As in Experiment 2, the first and second price estimates for the seed vehicle were removed (because the second estimates for the seed vehicle are simply copied from the value displayed on the screen) from the following analyses. As expected, price recall accuracy, the magnitude of the estimation error and the accuracy of participants' rank order correlations for the first set of estimates are very comparable with the results of Experiments 1 and 2 (see Table 1). Fig. 4 illustrates the impact that the Lexus LX470 seed price had on the second set of estimates across the three experimental conditions (the means and standard deviations are reported in Table 5). It is apparent from these results that the estimates observed in Experiments 1 and 2 are not simply the result of the stimulus context, as the automobiles' valuations are very comparable across all three experiments. As in Experiments 1 and 2, and using the same random effects GLS regression (with the log of the price estimates as the dependent variable), the two-feature (brand status and product class) model fits the data well for both sets of price estimates in the third experiment (two-feature model first set of estimates $R^2 = 0.394$; full model first set of estimates $R^2 = 0.438$; two-feature model second set of estimates $R^2 = 0.533$; full model second set of estimates $R^2 = 0.575$). We apply the model to only the full set of data here as the seed is the same in each experimental condition and our predictors (class and brand status) are not both present in Conditions 1 and 3. The intercept and the coefficients are significant for both the first and the second set of estimates, $p < 0.001$. The fitted models, with coefficients expressed in dollars, are

First set of estimates: $Y_e = \$22{,}240 + \$7272\ (\text{CLASS}) + \$14{,}819\ (\text{STATUS})$.

Second set of estimates: $Y_e = \$25{,}917 + \$10{,}880\ (\text{CLASS}) + \$23{,}831\ (\text{STATUS})$.

As expected, the intercept and the coefficients for the second set of estimates all increase in response to the seed price. The split-seed effect is again evident in the post-seed price profiles (Fig. 4) and verified by the regression models. The seed fact affects vehicles that share both critical features with the seed vehicle more than it affects vehicles that share a single feature or vehicles that do not share either critical feature. For example, consider the distinct differences in the magnitude of the revisions made to price estimates between basic passenger cars ($4122), basic SUVs ($14,524), luxury sedans ($10,769), and luxury SUVs ($25,887 averaged over all three conditions) (see Table 5 and Fig. 4). As predicted by the FBI model, the Lexus (luxury) SUV seed price has a small effect on basic passenger cars (that share none of the critical features with the seed vehicle); a moderate effect on basic SUVs (which share one of the critical features: product class); a moderate effect on luxury sedans (which share one of the critical features: luxury brand status); and, an additive effect on luxury SUVs (which share both critical features with the seed vehicle). Moreover, the effect of the seed price on estimate revisions for other luxury SUVs is very closely approximated by the effect on basic SUVs plus the effect on luxury sedans. This pattern of results cannot be accounted for by any of the previously identified modes of numerical estimation.

Finally, these results are also consistent with Experiment 2 in that we see a general resetting of the metric acting in concert with the FBI process (see Table 5 and Fig. 4). This effect is captured in the intercept of our model (i.e., α). Here, even though the increase in the mean price of basic passenger cars is small compared with the increase in luxury SUV prices (see Table 5 and Fig. 4), the difference is significant ($M = \$4122$; $t_{(1,146)} = -3.534$, $p < 0.001$) and consistent with a general resetting of the overall metric.

7. Discussion

We have argued that estimates of automobile prices are being inferred from knowledge of two critical features, which implies that the general pattern of price estimates we have seen in the first two experiments should not be context-specific. The results of Experiment 3 provide strong support for this prediction. Interestingly, in this experiment, we see a substantial upwards revision of the coefficient for brand status in the second set of estimates. Clearly, the precise weight of the coefficients will depend on the difference between the seed price and the previous estimates; therefore, the FBI model does not predict the exact amount of

Table 5

Experiment 3: mean estimates and standard deviations by condition for the first and second estimates (the first and second estimates for the seed vehicle, the Lexus LX470 (actual base price $99,950), have been removed from the data).

		First set of estimates		Second set of estimates	
		Mean	Standard deviation	Mean	Standard deviation
Condition 1	Luxury sedans	$41,762	$13,666	$52,531	$20,836
	Luxury sports cars	$43,986	$15,209	$55,571	$22,957
	Luxury SUVs	$50,048	$15,015	$74,452	$24,186
Condition 2	Basic passenger cars	$21,551	$12,060	$25,673	$17,474
	Luxury SUVs	$55,500	$24,698	$79,310	$39,721
Condition 3	Basic SUVs	$36,952	$15,442	$51,476	$22,876
	Luxury SUVs	$53,714	$22,750	$83,159	$25,378

the change in the price estimates as a result of a seed price. However, it does predict the direction of the shift (i.e., the second estimates shift up towards the seed price), and the general magnitude of the shift (i.e., the total amount of the shift is less than the difference between the seed price and the first price estimate).

8. General discussion

Overall, the evidence presented in this paper provides strong support for the FBI model of numerical estimation, which is able to both predict and explain the *split-seed effect* – a significant contribution to the extant literature as no previous theory of numerical estimation is able to account for this effect. The theory and results reported in this paper also point to some potentially fruitful areas for future research. For example, the present work focuses on one domain (i.e., automobile prices). Although automobiles are an important domain of real-world knowledge that, given the economic importance of the industry is interesting in its own right, we believe that this model should be applicable to other domains as well. Specifically, we expect that an FBI process will be adopted in domains that are structured by a small number of critical features. In such cases, knowledge of the domain is also likely to be feature based and estimates should reflect this structure.

As previously discussed consumer products are one domain that is explicitly, and intentionally, organized by features. Our results indicate that for automobiles two of these features are highly diagnostic price cues. In other domains, a different number of features may be more appropriate and in some cases these features may be best described by more than two levels (i.e., they may not be binary). We suspect that price knowledge is also feature based in many other product categories. Beyond price knowledge, features may also play an important role in the estimation process in a variety of other domains such as ambient temperature (e.g., latitude and season), household income (e.g., neighborhood and occupation), age of a tree (e.g., height and width), and so on.

However, consistent with prior research on numerical estimation (Brown, 2002; Brown & Siegler, 1993; Friedman & Brown, 2000a, 2000b) we also expect there will be important domain-specific differences in the selection of key features and the reliance on knowledge of the relevant range. Although we find that two features are used to generate automobile price estimates, it may be that in many domains one feature is sufficient. In the study of subjective geography (Friedman & Brown, 2000a, 2000b; Friedman & Montello, 2006), for example, being able to place a city into a regional category may be all that is required to estimate latitude. Similarly, Gigerenzer, Hoffrage, and Kleinbölting (1991) have argued that judgments of city populations can be made on the basis of a single cue. On the other hand, there may be domains in which it is common to use multiple features (e.g., event dating) or in which different people use a different number of features (e.g., novices versus experts). It would also be interesting to further examine the impact that a seed fact has general metric knowledge when an FBI process is being used to generate estimates. In any case, extending the current model to other domains is likely to be practically and theoretically informative.

From an applied perspective, the FBI model suggests that when consumers update their knowledge of the price of a particular product that revision will have implications for other products' prices to the extent that the revision affects the weights given to the critical features within that domain. This may explain why luxury companies like Mercedes have difficulty selling lower priced automobiles (Heinrich, 2003). People looking for a less expensive car are likely to estimate that a Mercedes is outside of their price range and, therefore, not include it in their consideration set. In addition, our results are consistent with the idea that short-term price promotions can affect a brand's image and consumers' price expectations over the long run (Greenleaf, 1995). For example, consistently putting a product on sale at a lower price tier may eventually convince consumers that it is a lower status product. As a result, they will be less willing to pay as much for it as they would for a higher status product.

Acknowledgements

The authors gratefully acknowledge the support provided by the Social Sciences and Humanities Research Council of Canada to the first author, as well as an NSERC operating grant awarded to the second author. The authors wish to thank Adam Finn, Richard Johnson, Kersi Antia, Dave Jobson and Gerald Häubl for their valuable comments.

Appendix A

Experiment 3 – Automobiles to be estimated by condition

Condition	Brand	Category	Model	True base price
1	Infiniti	Luxury SUV	QX45	$48,800.00
1	Acura	Luxury SUV	MDX	$50,300.00
1	Lexus	Luxury SUV	LX470	$99,950.00
1	Cadillac	Luxury SUV	Excalade	$76,530.00
1	Mercedes	Luxury SUV	M-Class	$51,100.00
1	BMW	Luxury SUV	X5	$58,400.00
1	Lincoln	Luxury SUV	Navigator	$72,125.00
1	Infiniti	Sp sedan	I35	$41,200.00
1	Acura	Sp sedan	TL	$40,800.00
1	Lexus	Sp sedan	IS300	$37,775.00
1	Cadillac	Sp sedan	Seville	$63,400.00
1	Mercedes	Sp sedan	SL Class	$127,500.00
1	BMW	Sp sedan	M3	$73,850.00
1	Lincoln	Sp sedan	LS	$43,750.00
1	Infiniti	Luxury sedan	M45	$62,000.00

Appendix A (*continued*)

Condition	Brand	Category	Model	True base price
1	Acura	Luxury sedan	RL	$55,800.00
1	Lexus	Luxury sedan	GS	$61,700.00
1	Cadillac	Luxury sedan	Deville	$55,685.00
1	Mercedes	Luxury sedan	E-Class	$71,350.00
1	BMW	Luxury sedan	5 Series	$66,400.00
1	Lincoln	Luxury sedan	TownCar	$57,345.00
2	Infiniti	Luxury SUV	QX45	$48,800.00
2	Acura	Luxury SUV	MDX	$50,300.00
2	Lexus	Luxury SUV	LX470	$99,950.00
2	Cadillac	Luxury SUV	Excalade	$76,530.00
2	Mercedes	Luxury SUV	M-Class	$51,100.00
2	BMW	Luxury SUV	X5	$58,400.00
2	Lincoln	Luxury SUV	Navigator	$72,125.00
2	Nissan	Passenger car	Sentra	$15,598.00
2	Honda	Passenger car	Civic	$16,100.00
2	Toyota	Passenger car	Echo	$12,995.00
2	Pontiac	Passenger car	Sunfire	$15,995.00
2	Dodge	Passenger car	SX 2.0	$15,195.00
2	Ford	Passenger car	Focus	$16,475.00
2	Hyundai	Passenger car	Accent	$12,895.00
3	Infiniti	Luxury SUV	QX45	$48,800.00
3	Acura	Luxury SUV	MDX	$50,300.00
3	Lexus	Luxury SUV	LX470	$99,950.00
3	Cadillac	Luxury SUV	Excalade	$76,530.00
3	Mercedes	Luxury SUV	M-Class	$51,100.00
3	BMW	Luxury SUV	X5	$58,400.00
3	Lincoln	Luxury SUV	Navigator	$72,125.00
3	Nissan	Basic SUV	Murano	$37,770.00
3	Honda	Basic SUV	Pilot	$41,000.00
3	Toyota	Basic SUV	4Runner	$39,220.00
3	Pontiac	Basic SUV	Aztec	$27,970.00
3	Dodge	Basic SUV	Durango	$41,975.00
3	Ford	Basic SUV	Explorer	$38,600.00
3	Hyundai	Basic SUV	Santa Fe	$22,595.00

References

Adelman, L., Stewart, T. R., & Hammond, K. R. (1975). A case history of the application of social judgment theory to policy formation. *Policy Sciences, 6*(2), 137–159.

Anderson, S. P., de Palma, A., & Thisse, J. (1992). *A discrete choice theory of product differentiation*. Cambridge: The MIT Press.

Becker, G. (1993). The economic way of looking at behavior. *Journal of Political Economy, 101*, 385–409.

Briesch, R. A., Krishnamuthi, L., Mazumdar, T., & Raj, S. P. (1997). A comparative analysis of reference price models. *Journal of Consumer Research, 24*(Sept), 202–214.

Brown, N. R. (1990). Organization of public events in long-term memory. *Journal of Experimental Psychology: General, 119*(3), 297–314.

Brown, N. R. (2002). Real-world estimation: Estimation modes and seeding effects. *The Psychology of Learning and Motivation, 41*, 321–359.

Brown, N. R., & Siegler, R. S. (1993). Metrics and mappings: A framework for understanding real-world quantitative estimation. *Psychological Review, 100*, 511–534.

Brown, N. R., & Siegler, R. S. (2001). Seeds aren't anchors. *Memory and Cognition*, 29.

Brown, N. R., & Sinclair, R. C. (1999). Estimating number of lifetime sexual partners: Men and women do it differently. *Journal of Sex Research, 36*, 292–297.

Brunswik, E. (1952). *The conceptual framework of psychology*. Chicago: University of Chicago Press.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments* (2nd ed.). Berkeley: University of California Press.

Burt, C. D. (1992). Retrieval characteristics of autobiographical memories. *Applied Cognitive Psychology, 6*, 389–404.

Conover, J. N. (1986). The accuracy of price knowledge: Issues in research methodology. *Advances in Consumer Research, 13*, 589–593.

Conrad, F., Brown, N. R., & Cashman, E. (1998). Strategies for estimating behavioral frequency in survey interviews. *Memory, 6*, 339–366.

Dehaene, S., & Marques, F. J. (2002). Cognitive euroscience. Scalar variability in price estimation and the cognitive consequences of switching to the euro. *The Quarterly Journal of Experimental Psychology, 55A*(3), 705–731.

Dhami, M. K. (2003). Psychological models of professional decision making. *Psychological Science, 14*(2), 175–180.

Doherty, M. E., & Brehmer, B. (1997). The paramorphic representation of clinical judgment: A thirty-year retrospective. In W. M. Goldstein & R. M. Hogarth (Eds.), *Research on judgment and decision making: Currents, connections, and controversies* (pp. 537–551). Cambridge, UK: Cambridge University Press.

Doherty, M. E., & Kurz, E. M. (1996). Social judgement theory. *Thinking and Reasoning, 2*(2/3), 109–140.

Estelami, H. (1998). The price is right... or is it? Demographic and category effects on consumer price knowledge. *Journal of Product and Brand Management, 7*(25), 4–66.

Friedman, W. J. (1993). Memory for the time of past events. *Psychological Bulletin, 113*, 44–66.

Friedman, W. J. (2004). Time in autobiographical memory. *Social Cognition, 22*, 591–605.

Friedman, A., & Brown, N. R. (2000a). Reasoning about geography. *Journal of Experimental Psychology: General, 129*(2), 193–219.

Friedman, A., & Brown, N. R. (2000b). Updating geographical knowledge: Principles of coherence and inertia. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26*(4), 900–914.

Friedman, A., & Montello, D. R. (2006). Global-scale location and distance estimates: Common representations and strategies in absolute and relative judgments. *Journal of Experimental Psychology: Learning, Memory and Cognition, 32*(3), 333–346.

Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review, 98*(4), 506–528.
Greenleaf, E. A. (1995). The impact of reference price effects on the profitability of price promotions. *Marketing Science, 14*(1), 82–104.
Grewal, D., & Marmorstein, H. (1994). Market price variation, perceived price variation, and consumers' price search decisions for durable goods. *The Journal of Consumer Research, 21*(3), 453–460.
Hammond, K. R. (1955). Probabilistic functionalism and the clinical method. *Psychological Review, 62*, 255–262.
Hammond, K. R. (1986). *A theoretically based review of theory and research in judgment and decision making: Theories, methods and procedures*. New York: Praeger.
Hammond, K. R., Stewart, T. R., Brehmer, B., & Steinman, D. O. (1975). Social judgment theory. In M. F. Kaplan & S. Schwartz (Eds.), *Human judgment and decision processes* (pp. 271–312). New York: Academic.
Heald, J. E. (1991). Social judgement theory: Applications to educational decision making. *Educational Administration Quarterly, 27*, 343–357.
Heinrich, S. (2003). Mercedes recasts brand as affordable: New Ad campaign. *National Post*, May 22nd, FP6.
Kalyanaram, G., & Little, J. D. C. (1994). An empirical analysis of latitude of price acceptance in consumer package goods. *Journal of Consumer Research, 21*(Dec), 408–418.
Lawson, R., & Bhagat, P. (2002). The role of price knowledge in consumer product knowledge structures. *Psychology and Marketing, 19*(6), 551.
Leavitt, H. J. (1954). A note on some experimental findings about the meaning of price. *Journal of Business, 27*(July), 205–210.
Levitt, T. (1980). Marketing success through differentiation of anything. *Harvard Business Review, 58*(January–February), 83–91.
Lucas, R. (1987). Adaptive behavior and economic theory. In R. Hogarth & M. Reder (Eds.), *Rational choice: The contrast between economics and psychology*. Chicago: University of Chicago Press.
Marques, J. F., & Dehaene, S. (2004). Developing intuition for prices in euros: Rescaling or relearning prices? *Journal of Experimental Psychology: Applied, 10*(3), 148–155.
Marshall, A. (1890). *Principles of economics*. London: MacMillan.
McFadden, D. (1999). Rationality for economists. *Journal of Risk and Uncertainty, 19*(1–3), 73–105.
Menon, G. (1993). The effects of accessibility of information in memory on judgements of behavioural frequencies. *Journal of Consumer Research, 20*, 431–440.
Monroe, K. B., & Lee, A. Y. (1999). Remembering versus knowing: Issues in buyers' processing of price information. *Journal of the Academy of Marketing Science, 27*(2), 207–225.
Nickerson, R. (1980). Motivated retrieval from archival memory. In J. H. Flowers (Ed.), *Nebraska symposium on motivation 1980* (pp. 73–119). Lincoln: University of Nebraska Press.
Paulos, J. A. (1990). *Innumeracy: Mathematical illiteracy and its consequences*. New York: Vantage.
Payne, J. W., Bettman, J. R., & Johnson, E. J. (1992). Behavioral decision research: A constructive processing perspective. *Annual Review of Psychology, 43*, 87–131.
Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. New York, NY: Cambridge University Press.
Putler, D. S. (1992). Incorporating reference price effects into a theory of consumer choice. *Marketing Science, 11*(3), 287–309.
Rao, A. R., & Monroe, K. B. (1989). The effect of price, brand name, and store name on buyers' perceptions of product quality: An integrative review. *Journal of Marketing Research, 26*(3), 351–357.
Shum, M. S. (1998). The role of temporal landmarks in autobiographical memory processes. *Psychological Bulletin, 124*, 423–442.
Thaler, R. (1985). Mental accounting and consumer choice. *Marketing Science, 4*(3), 199–214.
Von Helversen, B., & Rieskamp, J. (2008). The mapping model: A cognitive theory of quantitative estimation. *Journal of Experimental Psychology: General, 137*(1), 73–96.
Wallbaum, S. D. (1997). Seeding the fast food knowledge base: Long-term effects. *Poster presented at the 38th annual meeting of the psychonomic society, Philadelphia, PA*.
Wigton, R. S. (1996). Social judgement theory and medical decision making. *Thinking and Reasoning, 2*(2/3), 175–190.
Zacharakis, A. L., & Meyer, G. D. (1998). A lack of insight: Do venture capitalists really understand their own decision process? *Journal of Business Venturing, 13*, 57–76.