The authors integrate previous research that has investigated experimentally the influence of price, brand name, and/or store name on buyers' evaluations of product quality. The meta-analysis suggests that, for consumer products, the relationships between price and perceived quality and between brand name and perceived quality are positive and statistically significant. However, the positive effect of store name on perceived quality is small and not statistically significant. Further, the type of experimental design and the strength of the price manipulation are shown to significantly influence the observed effect of price on perceived quality.

The Effect of Price, Brand Name, and Store Name on Buyers' Perceptions of Product Quality: An Integrative Review

Since Leavitt (1954) examined buyers' tendencies to use price as an indicator of quality, numerous studies have examined the price-quality relationship with little consensus as to its magnitude, generalizability, or statistical significance. Various qualitative reviews of the evidence conclude that there seems to be a positive relationship between price and perceptions of quality for certain products and within certain price ranges; however, these effects are moderated by methodological deficiencies, a variety of contextual and situational factors, and a weak underlying theoretical explanation (Monroe 1973, 1977; Monroe and Dodds 1988; Monroe and Krishnan 1985; Olson 1977; Peterson and Wilson 1985; Rao and Monroe 1988; Zeithaml 1988).

The tradition in this research stream has evolved from considering the effect of only price on buyers' perceptions of product quality to including other intrinsic and extrinsic cues. To begin to unravel some issues and address some questions in the area, we integrate previous research that has examined the effects of price, brand name, and store name on product quality perceptions. In addition, we examine certain methodological differences among price—perceived quality research studies to determine whether those differences are associated with variations in outcomes among the studies. An examination of the criticisms in previous commentaries on the research stream generated hypotheses about four methodological variables to explain outcome variation across studies. We test these hypotheses using meta-analytic procedures on the results of previous research studies.

OVERVIEW OF THE RESEARCH DOMAIN

When suggesting that people may judge quality by price, Scitovsky (1945) pointed out that such behavior is not irrational; it simply reflects a belief that the forces of supply and demand would lead to a "natural" ordering of products on a price scale, leading to a strong positive relationship between price and product quality. Empirical attempts to verify this strong actual correlation have concluded that, generally, there is a positive correlation; however, though statistically significant, this correlation varies considerably across product-markets, producing an average correlation of $r = .27$ (Tellis and Wernerfelt 1987).

We examine another research stream that has attempted to determine whether buyers perceive a positive price-quality relationship. The statistical results across previous studies have been inconsistent and the com-
mentaries on past research have been rather critical. Possible methodological explanations for the inconsistent statistical results include the type of product, experimental procedures used, and uncontrolled or unmeasured individual response variations (Monroe 1977; Olson 1977). Focusing on the experimental procedures used produces four methodological issues that may explain why previous studies have generated such inconsistent results: number of cues, experimental design, price manipulation, and price level.

**Number of Cues**

Olson (1973) argued that consumers use a variety of cues to infer product quality. These cues include extrinsic cues that are not related directly to product performance and intrinsic cues that are derived directly from the physical product and, if changed, would change the product itself. For example, price, brand name, and store name (extrinsic cues) as well as nutrition content (intrinsic cue) would be used by consumers in their assessment of the quality of a breakfast cereal. Therefore, because numerous cues affect quality perceptions, the use of multiple independent variables in addition to the price cue is necessary for valid empirical tests.

Further, single-cue price-quality studies have been considered somewhat artificial; in the absence of any other information, subjects should naturally exhibit a positive price-quality effect (Enis and Stafford 1969). Therefore, the implicit inference is that single-cue price-perceived quality studies should manifest larger effects than multicue studies.

**Experimental Design**

Conceptually, it has been argued that the hypothesis of people using price as an indicator of quality is studied best by an idiothetic approach—that is, repeated measures over several prices—as the experimental situation created is analogous to a real-world situation of examining several different choices at different prices (Monroe and Dodds 1988). However, the use of a repeated-measures design (within subjects) has been criticized as being potentially artificial as subjects responding to several prices sequentially may guess the true intent of the researcher and respond accordingly (Sawyer 1975). Further, because variability due to individual difference is controlled in a within-subjects design, thus reducing error variance, within-subjects designs are likely to produce larger effects than between-subjects designs. For both of these reasons, price-perceived quality studies having within-subjects designs should show larger effects than studies having between-subjects designs.

**Price Manipulation**

In experimental research, generally the strength of the independent variable manipulation has an important bearing on whether statistically significant differences in the dependent variable occur. Price perception research suggests that people's perceptions of relative and not absolute differences in prices lead to differential responses (Monroe 1973; Monroe and Petroshius 1981). Therefore, we expect that the greater the relative difference between the prices used as levels of an independent variable in a given study, the larger will be the observed effect on subjects' perceptions of quality, and, ceteris paribus, the better the chance of detecting a statistically significant effect.

**Price Level**

Finally, it has been suggested that people are more likely to use price as an indicator of quality for relatively expensive products (Olson 1977). As the price level increases, the risk of an incorrect assessment increases and buyers often are less familiar with the product because of the infrequency of purchases. In such situations, simple learned heuristics based on folk wisdom such as "you get what you pay for" are likely to be used.

In sum, the conventional wisdom in the literature suggests that the magnitude of the price-perceived quality effect is influenced by at least four methodological factors: number of cues used, type of design, size of the price manipulation, and relative price level of the stimulus. Several other substantive factors have been proposed to influence the magnitude of the price–perceived quality effect (Monroe and Dodds 1988; Peterson and Wilson 1985; Zeithaml 1988), but our review focuses on the four methodological criticisms of previous price–perceived quality studies.

**Hypotheses**

The preceding commentary suggests four hypotheses about the impact of methodological variables on the observed price–perceived quality effect.

The price–perceived quality effect size will:

- **H₁:** vary positively with the price level of the test products.
- **H₂:** vary positively with the strength of the price manipulation.
- **H₃:** be smaller for multicue studies than for single-cue studies, and
- **H₄:** be smaller for between-subjects designs than for within-subjects designs.

**REVIEW PROCEDURES**

To find relevant research studies, we examined the references of previous review articles and conducted a computer bibliographic search. Because the hypotheses pertain to methodological variations across laboratory experiments, only laboratory studies involving one or more of price, brand, and store cues were used in the review. Further, when the results of a study were reported in multiple outlets, the study was included only once. A final set of 36 studies that collectively report 85 effects of price, brand name, or store name on perceptions of quality were used in the analysis and tests.

We computed the effect size index \( \eta^2 \) for all main effects by using
BUYERS' PERCEPTIONS OF PRODUCT QUALITY

(1) \[ \eta^2 = \frac{F(df_s)}{F(df_s) + df_{error}} \]

where \( F \) is the ratio reported from relevant ANOVA summary tables, \( df_s \) is degrees of freedom associated with the treatment effect, and \( df_{error} \) is degrees of freedom associated with the error term. Further, \( r^2 \) is a special case of \( \eta^2 \) when the \( F \)-ratio has only one degree of freedom associated with the numerator and is an effect size indicator for studies comparing only two groups. Cohen's (1977) \( d \) effect size indicator was used for one study and converted to \( r^2 \) by using

(2) \[ r^2 = \frac{d^2}{(d^2 + 1/p \cdot q)} \]

where \( p \) and \( q \) are the proportions of the sample that each mean represents (Rosenthal 1984). For studies using MANOVA the effect size used was (Bray and Maxwell 1984)

(3) \[ \eta^2(mult) = \text{Pillai's trace } / s, \]

where Pillai's trace is \( \Sigma \lambda_i / (1 + \lambda_i) \), \( \lambda_i \) is the eigenvalue of the \( i^{th} \) variable, and \( s \) is \( \min(\text{number of variables, hypothesis degrees of freedom}) \).

The design of the study (within or between) and number of cues used (single or multiple) were coded as 0, 1 dummy variables. Price level was determined by calculating the mean of the highest and lowest price treatments in each study \([\text{high} - \text{low}]/2\). The strength of the price manipulation was determined by computing a proportion variation between the highest and lowest price treatments for each product studied \([\text{high} - \text{low}]/\text{high}\) and ranged from 0 to 1.

The dependent variable measured in a study was determined by examining the indicators used. Perceived quality as the dependent variable was indicated by such terms as "durability," "reliability," "workmanship," "quality," "goodness," and "excellence."

RESULTS AND ANALYSIS

Description of the Data

The spectrum of products used as stimuli ranged from consumer nondurable products (e.g., butter, margarine) to capital equipment. However, for a preponderance of consumer goods, prices ranged from $.11 (table salt) to $400 (stereo receivers). Only four consumer products exceeded $100 in price and 60% cost less than $30. The strength of the price manipulations ranged from .18 to .86. Twelve of the 54 price-perceived quality results stemmed from a single-cue presentation and 19 results were obtained with a within-subjects design. Table 1 provides a stem and leaf plot, which is a data display technique analogous to a histogram (Rosenthal 1984), and a statistical summary of all price-perceived quality effect sizes. Table 2 provides a stem and leaf plot and a statistical summary of all brand- and store-perceived quality effect sizes. (A complete list of studies coded and their characteristics is available from the authors.)

Comparing Results Across Studies

Studies investigating the effect of price, brand, or store on buyers' product evaluations have tested whether there is a degree of association between these variables and buyers' evaluations. As these relationships have been studied several times, three questions should be addressed (Fleiss 1981):

1. Is there evidence that the degree of association is consistent across studies?
2. If consistent, is the common degree of association statistically significant?
3. What is the best estimate of the association, its standard error, and confidence interval?

Essentially, the first question seeks to determine whether the studies to be integrated have produced effect sizes that are from the same underlying population of effect sizes. A homogeneity of effect size test using an approximate chi square with \((K \text{ results } - 1)\) degrees of freedom addresses this issue. If the test fails to accept the null hypothesis of no significant differences, the conclusion is that the variation in results across the studies is not explained simply by sampling error and that systematic differences are present across the studies. Acceptance of the null hypothesis is evidence of homoge-

Table 1

<table>
<thead>
<tr>
<th>( \eta^2 ) Stem</th>
<th>Leaf</th>
<th>Summary statistics (all results)</th>
</tr>
</thead>
<tbody>
<tr>
<td>.8</td>
<td>7</td>
<td>Weighted mean, ( \eta^2 )</td>
</tr>
<tr>
<td>.7</td>
<td>0</td>
<td>Weighted standard deviation, ( s )</td>
</tr>
<tr>
<td>.6</td>
<td>3</td>
<td>Standard error</td>
</tr>
<tr>
<td>.5</td>
<td>0 0 5</td>
<td>95% confidence limit</td>
</tr>
<tr>
<td>.4</td>
<td></td>
<td>( \chi^2 ) (53)</td>
</tr>
<tr>
<td>.3</td>
<td>0 8</td>
<td></td>
</tr>
<tr>
<td>.2</td>
<td>0 0 1 3 5 5 6 8</td>
<td></td>
</tr>
<tr>
<td>.1</td>
<td>0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 2 2 3 4 5 6 6 6 6 6 7 7 8</td>
<td></td>
</tr>
<tr>
<td>.0</td>
<td>0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 2 2 3 4 5 6 6 6 6 7 7 8</td>
<td></td>
</tr>
</tbody>
</table>

\( K = 54 \)

*To read the plot, note that three effects are \( \geq .5 \) but \( < .6 : .50, .50, .55. \)
neity or consistency of the degree of association across studies and the results then can be integrated and analyzed. The statistic used to test the homogeneity of effect sizes was adapted from Pleiss (1981, p. 163):

\[ \chi^2 = \sum_{i=1}^{K} w_i (\eta_i - \bar{\eta})^2 \]

where:
- \(\eta_i\) is the individual effect size for each result,
- \(\bar{\eta}\) is the weighted estimator of effect size,
- \(w_i\) is the weight of each result, and
- \(K\) is the number of results being compared.

When results are compared across studies, substantial variation in average effect size can occur depending on whether individual effect sizes are weighted by sample size. Further, if sample sizes differ, the estimates of the larger sample studies will be more precise. Therefore, a weighting procedure that gives more weight to the more precise estimates is adapted from Pleiss (1981, p. 161):

\[ W_i = \frac{1}{\nu_i} \left( \sum_{i=1}^{K} (1/\nu_i) \right) \]

where:
- \(\nu_i = N_i/B_i + (\eta_i^2)/2N_i\),
- \(N_i = n_{i,E} + n_{i,C}\),
- \(B_i = (n_{i,E})(n_{i,C})\), and
- \(n_{i,E}\) and \(n_{i,C}\) are the sample sizes of the experimental and control groups, respectively.

This procedure minimizes the variances of the weighted estimates and produces an asymptotically efficient weighted estimator.

The homogeneity test for the price–perceived quality results using all 54 datapoints shows that the data are not homogeneous \(\chi^2(53) = 79.22\); Table 1), implying some systematic variation across these studies. The effect sizes for both the brand \(\chi^2(14) = 19.9\) and store \(\chi^2(16) = 6.13\) results are homogeneous; therefore these results can be integrated. As shown in Table 2, the association between brand and perceived quality is significant at the 95% confidence level and the best estimate of the association is the weighted mean of \(\bar{\eta} = .142\).

However, though the association between store and perceived quality is consistent, it is not significant at the 95% confidence level because the confidence interval for the weighted mean includes zero.

### Residual Analysis

The price–perceived quality dataset was examined for outliers and influential datapoints and a no-intercept regression model based on the hypotheses stated before was assessed for collinearity. Forty of the 54 results were used in this analysis as the other 14 results did not provide sufficient detail for coding of all four independent variables.

An examination of R-Student \(t_i\) values (an externally standardized residual) revealed four datapoints that were more than the conventional 1.5 standard deviations away from the centroid of the data cluster (Myers 1986). These four datapoints therefore were classified as outliers. Outliers are errors in the y direction caused by a large random error component in the response, producing a large residual. Datapoints also may be extreme in the x direction. Such datapoints tend to have a large influence on the regression, producing large hat diagonal \(h_{ii}\) values. Using the traditional cutoff of \(2p/n = .20\) for hat diagonal values (Myers 1986), we identified three data-
points as influential. These three datapoints represent relatively high priced industrial products and therefore have a large influence on the slope of the regression line. Thus, seven results were segregated for the meta-analysis, leaving 33 datapoints for the tests of the hypotheses.

For the reduced set of 33 price–perceived quality data­points used in the regression analysis, a homogeneity analysis indicated that the data were homogeneous ($\chi^2(32) = 37.64$). This dataset had weighted mean $\tilde{\eta}^2 = .117$, weighted $s = .147$, standard error $=.025$, and 95% confidence interval $\pm .049$. Thus, for this reduced set of 33 results, 3632 additional studies with a zero average effect size would be needed to reduce the statistical significance below the .05 level. Next, the 14 datapoints that had been eliminated for the regression analysis were retrieved and added to the 33 datapoints used in the regression analysis, producing a set of 47 homogeneous results ($\chi^2(46) = 49.73$) with weighted mean $\tilde{\eta}^2 = .116$, weighted $s = .125$, standard error $=.018$, and 95% confidence interval $\pm .035$. These summary results are very robust; an additional 8897 studies with an average zero effect size would be needed to reduce the statistical significance of the price–perceived quality relationship below the .05 level. This result lends confidence to the claim that the overall results are not due to sampling bias in the studies selected for review (Rosenthal 1984).

Multicollinearity Diagnostics

The possibility of linear dependencies in the regressor variables was investigated and a summary of the multicollinearity diagnostics is included in Table 3. The relatively low variance inflation factors are indicative of low multiple correlations among the independent variables. Further, the relatively high minimum eigenvalue ($\lambda_{\text{min}}$) and low condition number ($\phi = \lambda_{\text{max}}/\lambda_{\text{min}}$) confirm the absence of any significant multicollinearity. Thus, we have little concern about multicollinearity being a problem in this model.

Regression Analysis and Findings

To test the four hypotheses, we performed a multiple regression analysis on the dataset using the weighted price–perceived quality effect sizes as the dependent variable. The regression model tested the relationship between price–perceived quality effects and price level, size of price manipulation, number of cues, and research design. As is evident from Table 3, the model performed well. The preceding theoretical argument suggests a positive coefficient for price level and price manipulation and a negative coefficient for number of cues. Also, within-subjects studies were expected to produce larger effects than between-subjects studies. The results of the regression run are reported in Table 3.

Support for the hypotheses was determined by examining the sign and significance of the coefficients for the reduced dataset comprising the 33 homogeneous results. As shown in Table 3, price level is not significantly related to size of effect (H1). However, the relative strength of price manipulation is associated significantly and positively with effect size (H2). Though there is no statistically significant difference, multicue studies produced marginally larger effect sizes ($\tilde{\eta}^2 = .124$) than single-cue studies ($\tilde{\eta}^2 = .115$) (H3). Finally, within-subjects designs (repeated measures) generated larger effects ($\tilde{\eta}^2 = .20$) than between-subjects designs ($\tilde{\eta}^2 = .10$) (H4).

DISCUSSION

Single Versus Multiple Cues

Using criteria suggested by Cohen (1977), the price effect on perceived quality for consumer products (.12) is moderately large and statistically significant. The effect of brand name on perceived quality is slightly larger (.14) and statistically significant, but the effect of store name on perceived quality is small (.05) and not statistically significant. Moreover, though statistically not significant, multicue studies generated slightly larger effects (.124) than single-cue studies (.115). Taken together, these results suggest that a price–perceived quality main effect generated from a multicue study is not necessarily smaller than that generated from a single-cue study. This conclusion supports an earlier observation by Monroe and Krishnan (1985) that price–perceived quality effects actually increased in the presence of brand information. Thus, rather than brand name or other cues suppressing price as an indicator of product quality, a reinforcing effect is likely if the multiple cues are consistent in their signaling of quality. This finding is also consistent with the covariation literature (Nisbett and Ross 1980).

Price Manipulations

The regression analysis indicates that the relative strength of the price manipulations had a significant ef-

| Table 3: SUMMARY OF REGRESSION ANALYSIS: PRICE-PERCEIVED QUALITY EFFECTS (33 homogeneous results) |
|-------------------------------+------------------|
| Collinearity                  |                  |
| $VIF_{\text{max}}$            | 7.51             |
| $\lambda_{\text{min}}$        | 0.09             |
| $\phi$                        | 32.88            |
| Model                         |                  |
| $R^2$                         | .85              |
| $R^2_{\text{adj}}$            | .83              |
| PRESS                         | 1870.32          |
| Parameters                    |                  |
| Price level                   | $.10 (p < .11)   |
| Standardized estimate         | $.30             |
| Manipulation                  | 14.51 (p < .01)  |
| Standardized estimate         | 1.19             |
| Design                        | 4.69 (p < .08)   |
| Standardized estimate         | 1.17 (p < .66)   |
| Cues                          | 7.7              |
| Standardized estimate         | .19              |
fect on the observed price–perceived quality relationship. Methodologically, this finding is an important explanation for some of the inconsistency in statistical findings across previous price–perceived quality research and therefore is of significance in designing price perception research in the future; strong manipulations are more likely to produce large effects, thereby enhancing the chance of a statistically significant effect.

When buyers do infer a positive relationship between price and product quality, they are likely to compare the price of the product against another price (price in memory or price of an alternative option). If the actual price is perceived as significantly different from this reference price, the higher priced option is likely to be perceived as being of higher quality. The key point is that judgments of quality based on price information are necessarily comparative and perceived differences in prices lead to relative judgments that product quality varies significantly. In the research reviewed here, the greater the ratio between the highest and lowest experimental prices, the greater is the price–perceived quality association. Regardless of whether a between-subjects or within-subjects design is used, the greater the difference in the price treatments, the more likely individuals are to perceive differences in prices and begin to make quality inferences. Thus, weak price manipulations are a possible explanation for previous price–perceived quality studies’ failure to find a statistically significant price main effect.

Research Design

The finding that within-subjects designs generated significantly larger effects than between-subjects designs (.20 vs. .10) does not necessarily mean that use of such designs in price research is incorrect or artifactual. It does mean that researchers must be aware that within-subjects designs generally are more powerful and should consider this point when designing their studies and developing their conclusions. The choice of the appropriate design and price operationalization should be consistent with the particular research question being studied. For instance, Monroe and Dodds (1988) make the point that between-subjects designs are more appropriate for price present/absent studies that examine the marginal impact of price information.

Further, a managerial implication consistent with the reference price argument mentioned before is evident here. Multiple responses to different prices (as in walking down an aisle in a store) are a more powerful means of generating price-quality inferences than exposures to single prices (as in television commercials). For the latter case, the cognitive effort involved in retrieving a reference price and using it for comparative purposes may result in weaker effects.

Price Level

One obvious implication of the lack of a significant association between price level and size of effect is the dearth of studies using products that are relatively higher priced and less frequently purchased. Indeed, for four results for consumer products priced at more than $100, the weighted mean effect is $\eta^2 = .39$. Two of the results classified as outliers are for products priced above $100 with an average effect size of .57. (The other two outliers have effect sizes of .00.) Clearly, more research using moderately priced and relatively expensive products is necessary to establish whether the price–perceived quality relationship differs by relative price level and whether it differs by the general nature of the product.

CONCLUSIONS

Meta-analyses often have been criticized for two reasons. First, the “file drawer” problem prevents publication of many studies that do not show significant effects. Second, the notion of statistical significance is considered moot when a census of results is being considered. In our review, we attempt to address the first issue by estimating the number of nonsignificant (and thus “filed away”) studies required to negate the values computed from published results. The evidence suggests that the findings are relatively robust. Second, the notion of statistical significance is an important one, as the findings are being generalized to a population of effects. In other words, the studies considered represent a sample of effects of the phenomenon under consideration. The interested reader will find considerable information on this issue in books by Fleiss (1981), Hedges and Olkin (1985), Hunter, Schmidt, and Jackson (1982), and Rosenthal (1984).

Our review, in addition to providing estimates of the size of the effect of price, brand name, and store name on quality perceptions, investigates four possible explanations for the lack of consistent results across studies. Two methodological variables—type of research design and strength of price manipulations—provide some explanation for the inconsistent findings from previous price–perceived quality research. The review also indicates the relative narrowness of previous price–perceived quality research. At this point in time the phenomenon has been shown to be present at a moderate level of association for relatively lower priced, frequently purchased goods, but whether the strength of the association increases for higher priced, less frequently purchased goods has not been documented adequately. The quality perception research domain should be extended also to services and purchases made by business and institutional buyers.

Though our integrative review answers some questions about the effects of price, brand name, and store name on buyers’ perceptions of product quality, it also isolates important issues that warrant further investigation. Still needed is research on how quality perceptions are formed and how these quality perceptions influence perceptions of value, product or service benefits, and eventual choice. Further, the interactive effect of brand name, store name, and price on quality perceptions should be investigated. Finally, the effect of consumers’ prod-
uct knowledge may moderate the price–perceived quality effect.

REFERENCES


