A perennial question among applied marketing researchers is whether to measure stated or derived importance. The debate focuses on whether stated or derived importance is a better method, whether one is more valid or more actionable than the other and so on. Much of this attention is misguided, however, resulting from the mistaken conflating of two similar, but not identical, concepts. We illustrate an under-appreciated point made by Myers and Alpert over 30 years ago – that the choice between stated and derived importance is a false dichotomy: the two methods measure different constructs, they accomplish different objectives and they fulfill different information needs. Drawing upon brand studies with choice-based derived importance models and customer satisfaction studies with regression-based derived importance models, we show that, when done properly, both stated and derived methods have solid predictive validity, albeit with different strengths and weaknesses.

Motivation – Two Case Studies
Two disguised case studies illustrate what can happen if you measure importance badly.

#1 Suckered by Stated Importance
A service company wanted to know which aspects of its service most satisfied its customers. They asked 400 of their customers to rate the importance of each of the aspects on a scale from 0=Not Important At All to 10=Critically Important. When the results came back, all of the aspects had average importances in the range of 7.4 to 7.8. The survey didn’t give the service company any valuable feedback about which aspects of the service customers valued more than others, so it was a waste of a few tens of thousands of dollars and some goodwill, because some customers expected the service provider to make changes based on the survey results.

#2 Derived Importance Debacle
In the 1980s a medical supplies manufacturer had a 60% share of their market. A fancy consultant convinced the manufacturer that it should be using derived importance modeling to quantify the impact of attributes on customers’ choices. The consultant suggested a super-sophisticated method called multiple regression. He even put it in quotes, “multiple regression,” so that it would be clear to senior management how very cool and new and sophisticated it was to use regression instead of the stated importance methods the company had been using. So when the old stated importance methods said that ease of use was important to customers, the consultant and his regression analysis said that ease of use wasn’t important at all and that the key to incremental sales was size - the smaller the better. The manufacturer redirected new product development efforts away from easy to use products and toward small ones. The next year, a competitor launched an especially easy to use product and grew from a 15% share to 50%, almost overnight. The manufacturer, caught flat-footed, dropped from 60% to 30%, also almost overnight. Within a couple of years, the decision to emphasize size at the expense of ease of use had cost the manufacturer hundreds of millions of dollars. What happened?

These two case studies and many others like them illustrate some of the pitfalls associated with stated and derived importance measurement.

Introduction
Marketers often think of the products they sell as composites of discrete attributes. They assume consumers evaluate and decide among products on the basis of these attributes. As a result, the attributes feature prominently in marketing plans and in marketing communications like advertising, PR, promotions and packaging. Researchers assisting this marketing process study consumers’ beliefs about the attributes and they try to discern which attributes influence consumers’ decisions more than others. In seeking to understand the “importance” of the various product attributes, applied researchers typically distinguish between “stated” and “derived” importance measures.

In stated importance measurement the researcher elicits direct responses from respondents about attribute importance.
These might be ratings, rankings, constant sum scales, Qsorts or best-worst scaling, to name just a few methods.

Researchers using derived importance measurement rely on statistical modeling to reveal the influence of the various attributes. A statistical model wherein the researcher uses respondents’ rating of each of the product attributes to predict an outcome variable (an overall brand rating, a purchase likelihood, a brand choice) yields a coefficient (a weight) for each attribute; these coefficients become the derived measures of attribute importance.

**Standard Practical and Theoretical Considerations**

The efficiency of derived importance appeals to applied researchers. In terms of questionnaire real estate, derived importance can be had for free. Most questionnaires already capture attribute performance scores and overall outcome measures - all the raw materials statistical analyses for derived importance need. Asking respondents to supply stated importances, however, involves extra questions, which means time for the respondents and extra money from the research sponsors.

Both stated and derived importance face some well-known (and some less well-known) conceptual challenges to their validity. For one thing, researchers worry that stated importance measures may be prone to a pro-social bias. A pro-social bias occurs when respondents attach greater importance to attributes they perceive to be socially preferable in a survey environment than they do in real purchase decisions. For example, respondents may rate “green” attributes as more important than economic or performance attributes in a survey, but then discount those green attributes when they buy their Hummers. Or again, physicians may claim to prescribe the best drug for their patients’ care, when in practice they prescribe the brand whose sales reps visit most often and buy pizza for the office nurses.

Researchers often find respondents give all attributes high importance ratings. In the absence of constraints preventing them from doing otherwise, respondents find it easy, in fact sensible, to “want it all.”

Some of the strongest evidence against the validity of stated importances dates back to the early 1970s. At the time, a method called the Multi-Attribute Attitude Model was in vogue and it caught the attention of a number of researchers. Attributed to Fishbein (1967) and Rosenberg (1956), the Multi-Attribute Attitude Model suggests that one can approximate a person’s overall attitude toward an entity by adding together the products of each of the entity’s attributes multiplied by the respective importances of those attributes. So if a product has three attributes, price, accuracy and ease of use, the overall attitude would equal

\[
(\text{importance of price} \times \text{price performance score}) + (\text{importance of accuracy} \times \text{accuracy performance score}) + (\text{importance of ease of use} \times \text{ease of use performance score})
\]

This sounds pretty reasonable, and something like it must be true if the attention marketers pay to attributes is warranted.

Unfortunately for the Multi-Attribute Attribute Model, it performed poorly in tests of how well the computed overall attitude correlated with overall ratings or market shares. In fact, researchers found that the Multi-Attribute Attitude Model often performed worse when the attribute importance weights were included in the calculation than when they were left out – meaning that the stated importances essentially harmed predictive validity, a pretty damning finding indeed (Bass and Wilkie 1973; Wilkie and Pessemier 1973).

Not that derived importance enjoyed rosy prospects, either. Statistical models for derived importance operate off of the correlation each attribute has with the outcome variable, and this statistical modeling falls apart pretty quickly if the data violate some specific, and demanding, assumptions about those correlations. One type of miscreant correlation involves a statistical condition called multicollinearity. Multicollinearity means that high correlations occur among the attribute scores; unfortunately this happens pervasively in applied marketing research data sets. In 1920, a psychologist named Thorndyke noticed that when respondents like something, they tend to rate it more highly on all its attributes, while if they dislike it, they tend to rate its attributes lower across the board. He called this the “halo effect” and it often causes the high inter-correlations among all of a product’s attributes that constitute multicollinearity. Multicollinearity happens to a greater or
Stated “Versus” Derived Importance: A False Dichotomy
Taking a closer look at what these two methods really measure

lesser extent in almost all attribute surveys, and it produces a variety of nasty consequences, primarily because it causes very large estimation error around the coefficients (while invisible in a single analysis, this causes instability across studies or across segments). This large estimation error can lead to non-significant stat tests for attributes known to be influential, suggesting that they do not matter to consumers. If that were not bad enough, “reversals” can occur where attributes known to be important (e.g. price) can have smaller coefficients than attributes known to be less important – sometimes the coefficients can even be negative, suggesting that presumably valuable attributes are in fact harmful.

Another aspect of correlation can also cause trouble for derived importance. Because derived importance operates from the correlation between attributes and an outcome variable, if either is constant, the correlation will be zero by definition. So the less an attribute’s score varies across observations, the less able derived importance is to identify it as important. Famously, “must have” attributes having four wheels on a car or safety in an airline produce attributes whose ratings should not vary systematically across respondents (all cars have four wheels, and all airlines are about equally safe). As a result, derived importance will not identify a car’s having four wheels or an airline not crashing and killing its customers as important attributes.

While stated importance usually results in an importance measure for every observation/respondent, derived importance is a measure averaged across all observations used in the analysis.; Moreover, since derived importance models are being run cross-sectionally, or across respondents, it focuses on the differences between the extremes. In the case of regression, this means that the variance it seeks to predict is between respondents with higher scores on the outcome variable and those with lower scores. If these are customer satisfaction scores, for example, the analysis tells us how important each attribute is in differentiating between more satisfied and less satisfied customers. They can tell us how to make dissatisfied customers satisfied, but they cannot tell us how to make extremely satisfied customers more satisfied.

Finally, derived importance modeling, because it requires this cross-sectional statistical model, is sample size intensive. For average, well-behaved data the standard recommendation calls for 10 respondents for every variable in the model. So if we have 15 variables, that means 150 respondents. That’s for well-behaved data and typical survey data is notoriously ill-behaved, as noted above, so we may often need 20 or 30 respondents per variable. While this does not threaten all derived importance analyses, for some applications in some countries or sectors, samples of this size may not be attainable.

The following table summarizes the strengths and weaknesses of commonly used methods for stated and derived importance:

One wonders, with this witches’ brew of trouble for both stated and derived importance, why applied researchers still measure them so naively – and they do, with stated importance ratings and multiple regression derived importance analysis being both the norm among commercial marketing research firms and egregious examples of methods known to be flawed – even though better methods for doing stated and derived importance exist, as shown below.
Importance and Determinance
A researcher performing both stated and derived importance often finds the two produce disparate results. Sometimes researchers do silly things like plot stated and derived importance on Cartesian coordinate axes and make up stories about how some attributes are “hidden satisfiers” while others are “table stakes.” The reason this is a silly thing to do owes to a crucial distinction, made over 30 years ago by Myers and Alpert (1977), but that applied marketing researchers either never learned, or have forgotten. Myers and Alpert note that stated and derived importance measure different things. Using a Webster’s dictionary definition of importance “weighty, momentous, of great consequence, significance or value” they note that the definition fits stated importances. Derived importances, however, fit a different construct, which Myers and Alpert named “determinance.” To be determinant, an attribute must satisfy two conditions:

1. It must be important, in the sense defined above
2. It must vary systematically across outcomes (across observations in the case of regression models and across brands in the case of choice models).

Note that the second condition just describes one of the correlation problems we noted above for derived importance – the case where outcomes were not statistically related to an attribute score because the attribute had little or no systematic variance across respondents (or brands).

Stated importances satisfy only condition #1 above, so they measure importance. Derived importances satisfy both conditions 1 and 2, so they measure determinance. “Must have,” “table stakes” or “hygiene” attributes will be constants, so they will be important, but not determinant. This clear thinking, developed by a pair of marketing academics, should have prevented 30+ years of confusion over stated and derived importances. Unfortunately it did not.

So, an attribute that is high on derived importance is usually high on stated importance as well and also has a high variance in performance that correlates with the variance of the dependent variable. The only true disconnect between the two might occur when a variable is high on derived importance and low on stated importance.

Improving Stated and Derived Importance Measurement
It is easy to do both stated and derived importance badly, and badly is in fact the status quo, as noted above. A quick view of marketing research firms websites, white papers and point-of-view documents show that most of them collect stated importance ratings (which were proven to have zero or negative predictive validity over 35 years ago) and they calculate derived importances via correlation or regression models, which suffer from the correlation problems noted above.

Methodologically sound ways exist to collect valid stated and derived importance measures. Valid measures require more thought and methodological skill to produce, however. For example, valid stated importance methods result from forcing respondents to make tradeoffs among attributes. These range from a little more complex than simple rating scales (Q-sort, constant sum) to much more complex (best-worst scaling). These methods have been shown to outperform simple stated importance ratings so far as to resuscitate the otherwise defunct Multi-Attribute Attitude Model (Chrzan and Golovashkina 2005).

To estimate derived importance well, one must use statistical models that anticipate and accommodate multicollinearity. A class of models that use an “averaging over orderings” approach manage to accomplish this. Regression based models exist (Shapley 1953, Budescu 1993) but information theoretic models (Theil 1987) have further theoretical benefits, plus the practical advantage (Soofi, Retzer and Yasai-Ardekani 2000) of extending to cover both predictions of rating scale outcomes (averaging-over-orderings versions of regression and PLS) and categorical outcomes like brand choices (averaging-over-orderings versions of logistic regression and multinomial logit).

Which is Better, Stated or Derived Importance?
With the distinction from Myers and Alpert (1977) in mind, one can identify situations more appropriate for collecting (stated) importance and other situations more appropriated for determinance (derived importance).
Derived importance will be more appropriate to use when

a. The set of attributes and their ranges will remain stable in the period the research results will affect,
b. The study sponsors want to make improvements on current attributes within the range in which they currently vary and
c. The target population is relatively homogeneous, so the averaging effect of the analysis would not bias the resulting importances significantly.

Factors suggesting the use of stated importance include:

a. Small sample size available for the analyst precludes derived importance models
b. Study sponsor wants to know how new attributes, or how expanding the range to new levels of existing attributes, will affect outcomes
c. The study includes a large number of attributes (say more than 10 or so), a condition where even averaging-over-ordering statistical models may have trouble apportioning derived importances,
d. The importance of items differs widely for individual respondents (respondent heterogeneity), or
e. The detection of heterogeneity among respondents is desired because of further analysis (e.g., segmentation).

In terms of quality, a series of studies have shown that stated and derived importances, when conducted competently, have very similar levels of predictive validity, probably the most important measure of research quality. These studies have been published in two journal articles and they are summarized below.

Borrowing from the old Multi-Attribute Attitude Model literature, stated and derived importance measures populated equally successful predictive models for rating scale outcome variables (Chrzan and Golovashkina 2005), at least for stated measures forcing attribute tradeoffs. Using a large controlled experiment, Chrzan and Golovashkina (2005) tested the predictive validity of several importance measurement methods for predicting rating scale outcomes. The following table shows the correlation of each method’s predictions with actual overall attitudes:

<table>
<thead>
<tr>
<th>Method</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard regression</td>
<td>.56</td>
</tr>
<tr>
<td>Averaging-over-orderings regression</td>
<td>.58</td>
</tr>
<tr>
<td>Importance ratings</td>
<td>.44</td>
</tr>
<tr>
<td>Unbounded ratings</td>
<td>.50</td>
</tr>
<tr>
<td>Magnitude estimation</td>
<td>.58</td>
</tr>
<tr>
<td>Q-sort</td>
<td>.60</td>
</tr>
<tr>
<td>Constant sum scaling</td>
<td>.60</td>
</tr>
<tr>
<td>Maximum difference scaling</td>
<td>.64</td>
</tr>
</tbody>
</table>

Note that the three stated importance methods that force respondents to make tradeoffs – (Q-sort, constant sum scaling and maximum difference scaling) outperform stated measures that do not force tradeoffs. Of the two derived importance models, averaging-over-orderings regression outperforms standard regression in terms of predictive validity, confirming the results of earlier comparisons (Chrzan, Retzer and Busbice 2003).
Similarly, a series of 10 studies compared the ability of stated and derived models to predict brand choice. In these studies, the derived model was the MNL choice model while stated rank order importances were used to populate a lexicographic choice model (Chrzan and Malcom 2009). Stated importances outperformed derived importances in predicting brand choices – in all 10 studies - sometimes only slightly but always significantly:

<table>
<thead>
<tr>
<th>Study</th>
<th>n</th>
<th>% fit MNL</th>
<th>% fit Lexicographic</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pizza</td>
<td>402</td>
<td>.48</td>
<td>.52</td>
<td>4.24</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Retail petrol</td>
<td>1,063</td>
<td>.32</td>
<td>.53</td>
<td>7.89</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Smartphone (consumer)</td>
<td>502</td>
<td>.38</td>
<td>.46</td>
<td>5.89</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Smartphone (B2B)</td>
<td>502</td>
<td>.33</td>
<td>.40</td>
<td>4.67</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Oral contraceptives</td>
<td>200</td>
<td>.26</td>
<td>.31</td>
<td>2.13</td>
<td>.034</td>
</tr>
<tr>
<td>Investment services</td>
<td>1,602</td>
<td>.44</td>
<td>.53</td>
<td>11.87</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Athletic shoes</td>
<td>1,001</td>
<td>.62</td>
<td>.66</td>
<td>4.72</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Service provider (consumer)</td>
<td>502</td>
<td>.47</td>
<td>.59</td>
<td>8.31</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Service provider (B2B)</td>
<td>502</td>
<td>.45</td>
<td>.55</td>
<td>7.29</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Restaurant delivery</td>
<td>1,027</td>
<td>.58</td>
<td>.64</td>
<td>5.39</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Note that stated and derived importances have different advantages that help them perform well in terms of predictive validity. Stated importances, available at the individual respondent level, boost the predictive validity of the Multi-Attribute Attitude Model and of choice models by incorporating respondent heterogeneity. Derived importances, on the other hand, by focusing particularly on determinant attributes, put more weight on the attributes that determine outcomes.

**Conclusion**

Stated and derived importance measure distinct, if similar, concepts: stated importance measures importance while derived importance measures determinance. We should thus not expect the two to produce the same priority of attributes and we should not think one superior to the other if they do. A good practice might be to stop naming both measures “importance,” at all, and to more precisely name stated and derived “importance” and “determinance,” respectively.

In terms of prediction, both stated and derived importance have roughly equal validity, but this is true only if they are both done well. Regression models that do not account for multicollinearity, correlation models that do not possess composition rules, or stated importance models that do not force attribute tradeoffs will suffer diminished validity. Either stated or derived importance done well performs better than either done poorly.
References


