BEST PRACTICES IN REVIEW SUBDIMENSIONS AND FILTERS

PIXLEE | TurnTo

Tailoring review subdimensions and filters for the exact needs of your different product categories is a powerful way to increase the value that your shoppers – and you – get from product reviews.

Questions you'll want to ask as you design your subdimension strategy include:

WHAT TO COLLECT

What will be most useful to shoppers? What topics will benefit from the focus and structure of subdimensions? How can I use subdimensions to help shoppers find the reviews that will be most valuable to them? How many should I collect?

HOW TO DISPLAY

Which subdimensions should be aggregated in the review summary? Shown along with the individual reviews? Used to power filters? And how should I configure the filters?

INTERNAL USE

How do I make use of subdimensions internally for product and marketing insights, or perhaps for CRM enhancement? What additional data should I collect for these applications?

WHAT TO COLLECT

Subdimensions come in two general types, serving two different purposes:

Rating Subdimensions

These subdimensions relate to the item purchased and are tied to the product SKU. They ensure that feedback on topics likely to be of particular interest gets collected. They provide structure to those data to enable easy, clear summarization (typically by showing frequency counts or averages) and to enable use in filters.

Customer Profile Attributes

These subdimensions capture the characteristics of the people who wrote the reviews, so shoppers can determine the relevance of each review to them and can easily find reviews by "people like them". They are tied to the customer and, once captured, will "follow" the customer around when they submit new reviews.

Rating Subdimensions

The backbone of most reviews is the free text body. This captures the reviewer's sentiment in her own voice, and it enables the reviewer to touch on any topic – whatever seems most important to her.

But the free text body also has limitations. The topics a reviewer chooses to address may not be the ones of most interest to the shopper. And when there are a large number of reviews – more than a shopper is ever going to read – summarization becomes important, and unstructured text is hard to summarize.

That's where rating subdimensions come in. By asking specific questions, rating subdimensions ensure that topics likely to be of the greatest interest to shoppers get addressed. And by using structured data types, rating subdimensions enable summarization of reviewer feedback over large numbers of reviews. Structured data also helps shoppers compare reviewer sentiment between items.

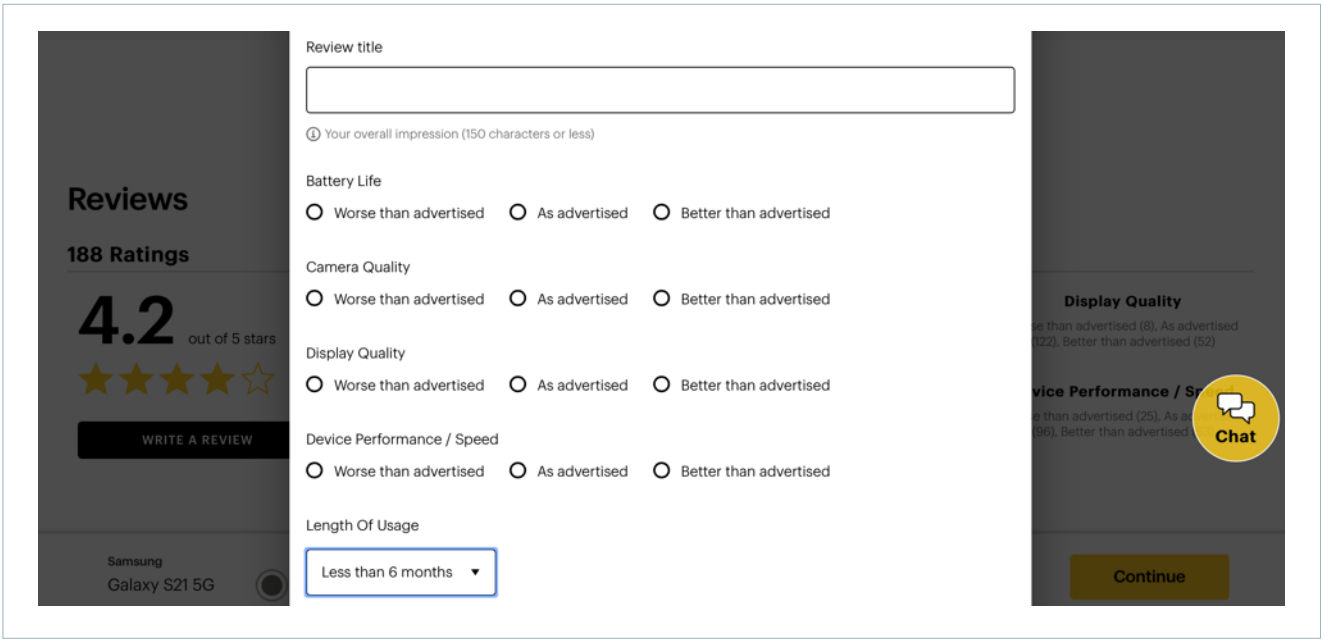
It's not easy to come up with just the right rating subdimensions for each category of product. Here are some best practices to guide you:

1 Compare real world experience to published specs.

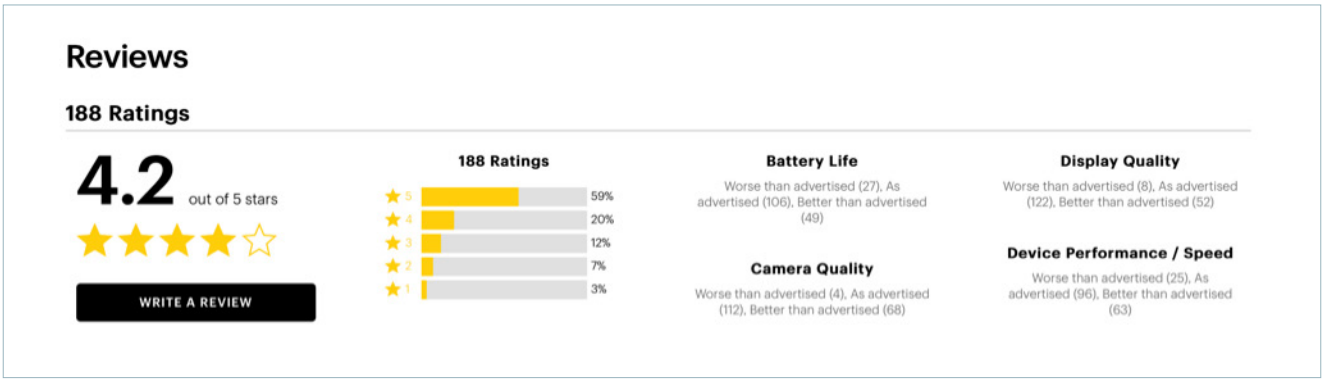
One of the most effective uses of rating subdimensions is to surface how the real world performance of products compares to the published specs. Typically, the data type for such a question is a spectrum with "matches the spec" in the middle and high/low variation on either side. Fit is the most widely used example of this. For apparel and footwear, it would not be helpful to ask "How big was it?" Obviously, bigger sizes are bigger than smaller sizes. So sub-dimensions for fit typically ask for fit relative to size with a spectrum of choices like "Runs Small – True to Size – Runs Large" or "Buy a Half Size Down – Buy Your Size – Buy a Half Size up". The same model can be used to collect feedback on many attributes of other types of products. If you sell mattresses which come in Firm, Medium, and Soft, instead of asking "How firm was it?", ask "How did the firmness compare to the rating?" If you sell battery-powered electronics, the official battery life is likely published. So instead of asking "How was the battery life?" ask "How long did the battery last compared to the spec?" Well crafted experience-vs-spec questions like these can provide information of solid gold to shoppers.

For summary display of these data, you have a couple choices. If you expect most responses to cluster together (ie most reviewers to agree with each other), show a bar with the published spec as the center point and high/low variation to the right and left. Then show the average response as a point on that range. If you think opinions will vary, use a distribution graph, or simply show counts, to show how many reviewers picked each option, so shoppers can see how much diversity of opinion there is. An average of "matches the spec" means something different if all reviewers chose that value than if half the reviewers chose a high value and half chose a low value.

As an example, Sprint has done a great job of using "As advertised" as the benchmark for rating subdimensions on the key performance attributes of the phones they sell. Customers all understand that the more expensive, higher-end phones are going to perform better. So asking about performance in absolute terms wouldn't reveal much. But by using this benchmark approach, Sprint enables shoppers to understand which phones are going to outperform or underperform their expectations, which helps shoppers make better choices and improves customer satisfaction with Sprint overall.



Sprint review collection form



Sprint review summary on product detail page

2 Quantify values whenever possible

When there isn't a published spec that a reviewer can use to benchmark their rating, subdimensions often default to a subjective score on a 5-point scale instead. For example, reviewers are often asked to "Rate the value" from low to high using 1 to 5 stars. Sometimes, subjective scales like these are all that's available, especially when the thing being rated is not well defined, like value. But whenever possible, we recommend trying to come up with an objective, quantified scale for the subdimension, so that the shopper reading the review knows exactly what the reviewer meant. For example, to create a subdimension for the durability of a running shoe, instead of asking "How durable was it?", ask "How many miles did you get before it showed wear?" For difficulty of assembly, instead of asking "How hard was it to assemble", ask "How many hours did assembly take?" For the durability of clothing in the wash, instead of asking "How did it hold up in the wash?" ask "How many wash cycles did you get before you noticed color fading?

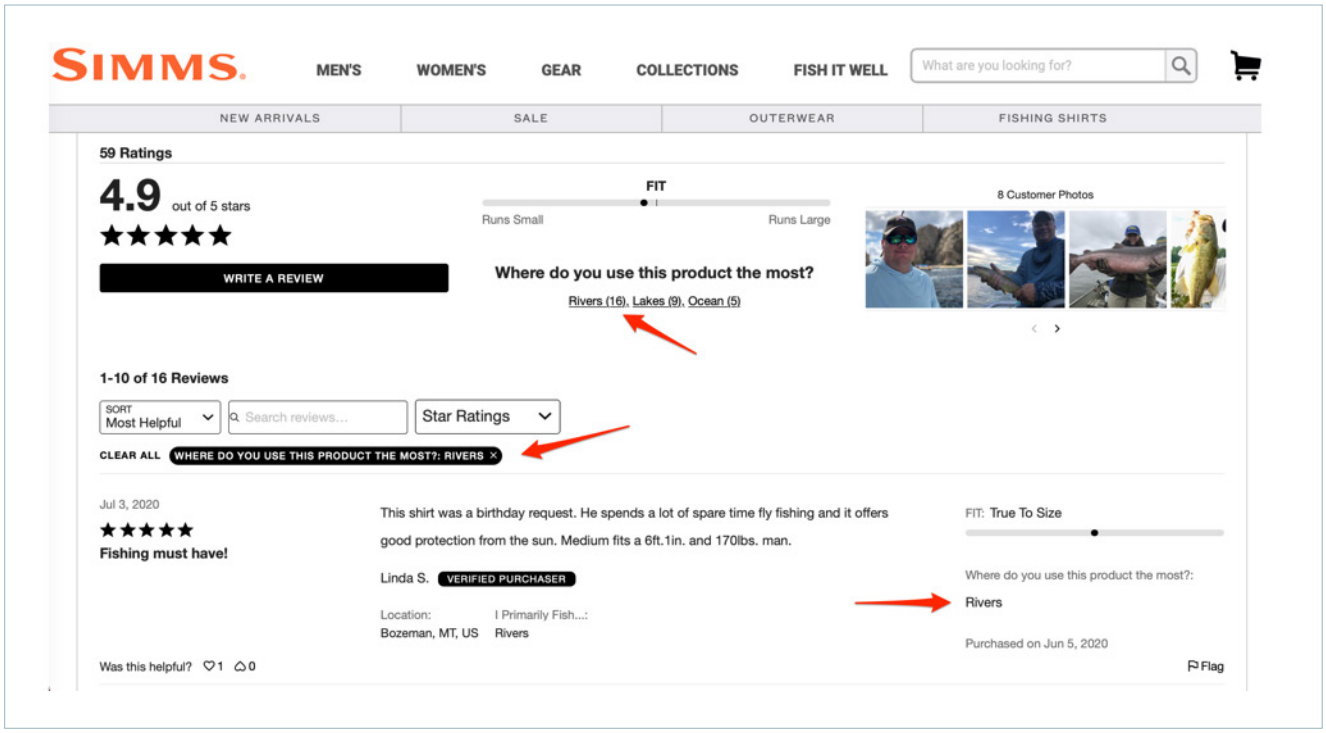
With quantified subdimensions, one challenge you'll face is determining the range of values you allow the reviewer to choose. You don't want to limit the choices for how long an assembly task takes to 5 hours maximum if some reviewers took a lot longer. An easy solution is to use a numeric value rather than a pick list. That approach works great if you just want to display an average. But as mentioned above, showing only an average can hide important information about how widely viewpoints vary. To show that, you'll want to display a distribution, which means you probably want to group the value choices into "buckets" and use a pick list. 1-2 hours, 3-4 hours… Then you have to figure out how many buckets to use. More buckets provide more granularity, but they can also make both collection and display too fussy, and you will end up with more empty buckets. In general, 5 to 7 are a good number of buckets. One technique, if you are not sure where the high and low extremes will fall, is to use open-ended greater-than or less-than buckets for the high and low ends of the scale. For example, <1 hour, 1-2 hours, 3-4 hours, >4 hours.

3 Use text picklists to create filters about use/purpose

While most rating subdimensions are designed to reveal information about the product, text picklist values are especially useful to help a shopper find reviews likely to be of particular interest based on how the item is used. For example, where an item has a number of possible uses, ask the reviewer "What was your primary purpose for buying this?" with a picklist of likely options. There are many variants on this, like, for outdoor gear, asking about the environment in which the item will be used; for shoes, asking the distance walked in them per day or the hours spent standing in them; or for dog toys, asking about the dog who will receive it (type, personality, preferences…). Then, use the answers to power a filter – either through a pull-down menu or through a clickable

item in a summary list. A shopper can use this filter to narrow a large body of reviews down to just those where the reviewer used the item for the same purpose that the shopper has in mind. For example, for a DSLR camera, ask whether the reviewer's primary use for the item is photographing kids, sports, nature, portraits, events… Then a shopper looking for a camera primarily for nature photography can easily see what others who used the camera in that way thought of it. For fishing gear, ask whether the primary location of use is lake, river, or ocean. This use of picklist values is a lot like customer profile attributes – see below – but the subdimension is tied to the item rather than to the reviewer. ie The reviewer might be primarily a lake fisherman (customer profile attribute), but he uses this particular item for salt water (rating subdimension).

An example of just that, on Simms Fishing:



A common use of picklists for rating subdimensions is to capture and display Pros and Cons. The reviewer is prompted with a list of options, and sometimes can add free text values. The results are sorted by frequency, implying that the more often chosen answers are the best or worst qualities of the item. While this use of subdimensions has appeal as a shortcut way to do sentiment analysis across a body of reviews, our user testing shows that shoppers often do not find value in this approach.

The prompt values are typically too generic or else are off base entirely. The free text responses are impractical to aggregate. And the frequency counts don't convey the same meaning that free text does. Compare the impact of text stating "The color range blew my mind" to "Pro: color (84)". The same tends to happen with another popular question, "Best uses", where the prompted values tend to produce obvious results.

4 Capture "variant purchased" to create variant filters, selectively

If your product pages group together reviews for variants of an item, it may be helpful for shoppers to be able to filter reviews to see just those by people who purchased the same variant they are considering. Variants typically include size and color, but some sites share reviews across items that differ in other ways, like packaging, material, style, or product configuration (like finish on furniture, feature bundles on equipment, components on electronics…). To enable filters based on item variants, you need to attach the variant purchased to the review. Ideally, that data is captured directly from the order, which saves a field on the review form. In practice, that level of integration can be challenging, and it's an OK alternative to simply include a variant-purchased subdimension on the review form.

To determine if variant filters would be valuable to shoppers, ask yourself if it's likely that reviews will be different across variants. Will people who bought large sizes possibly have different views than people who bought small sizes? Will people who bought the variant with the higher-end components have something different to say than those who bought the basic package? If so, include a rating subdimension for size purchased or component bundle purchased, and use that to power a variant filter. If not, save the clutter. You probably have an intuition about which types of variants will pass this test. If not, try including variant-purchased subdimensions for a period and have a look at the reviews you get for each. If there are no differences, you can retire that field.

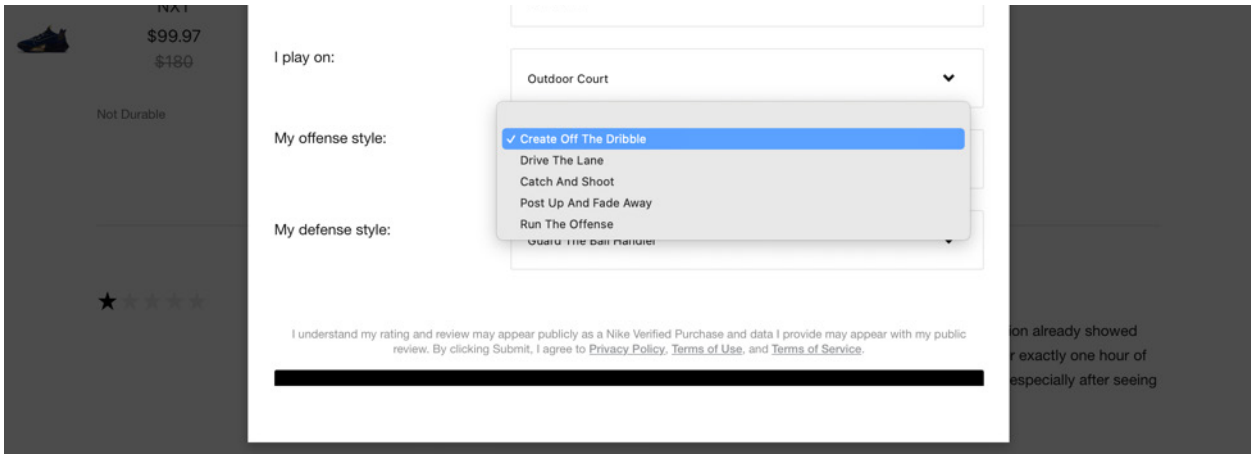
Customer profile attributes

Where rating subdimensions provide structured data tied to the product, customer profile attributes add structure to the description of the reviewer. The goal of customer profile attributes is to enable a shopper to determine whether the reviewer is similar or different from them in ways that would make the review more or less relevant.

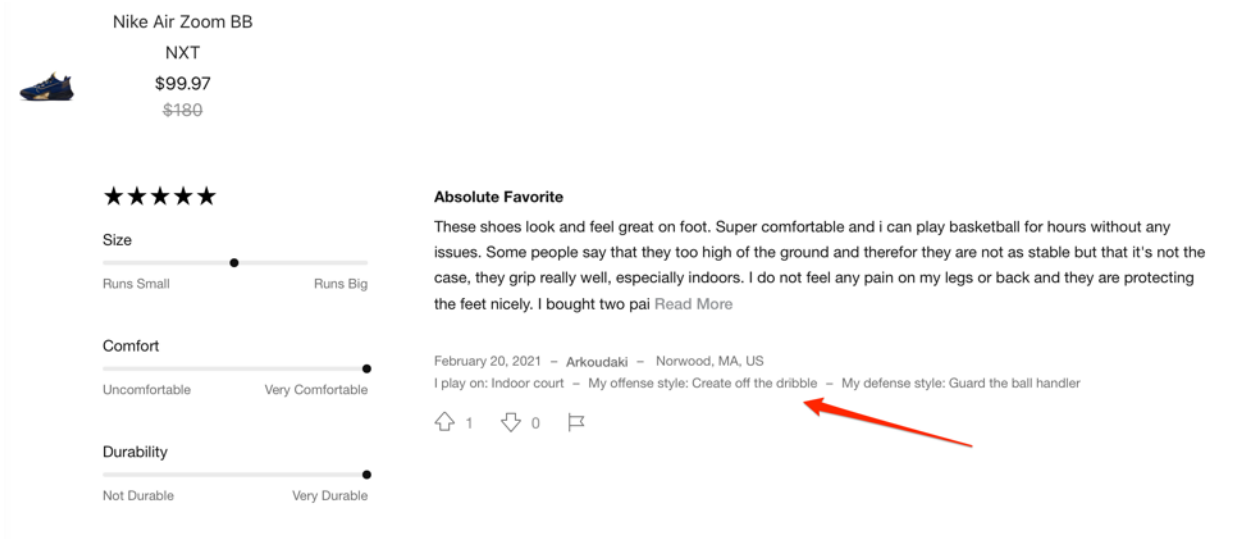
To design useful customer profile attributes, ask whether a shopper would find that information helpful in assessing the relevance of a review. Examples:

- For mattresses, ask about the reviewer's sleep position – front, side, or back
- For motorcycle helmets, ask the style of bike that the reviewer rides
- For products that require assembly or advanced skills to use, ask about the reviewer's level of expertise and experience
- For apparel, ask about the reviewer's body type
- For footwear, ask about the shape of the reviewer's foot
- For makeup or skin care ask about the reviewer's skin type and any special conditions

Nike does a brilliant job of making footwear reviews more useful by asking about the reviewer's play style, using different types of questions for different sports. For example, for basketball they ask:

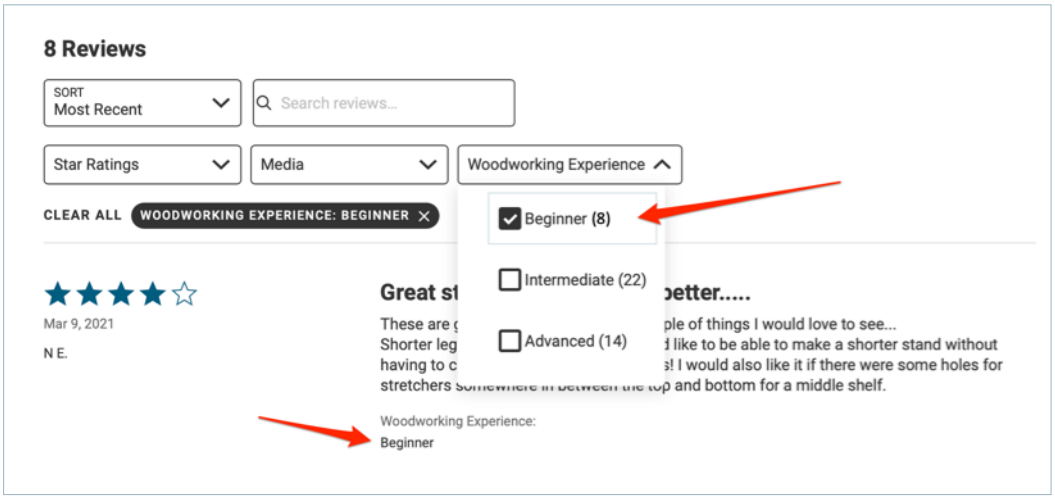


Nike review collection form



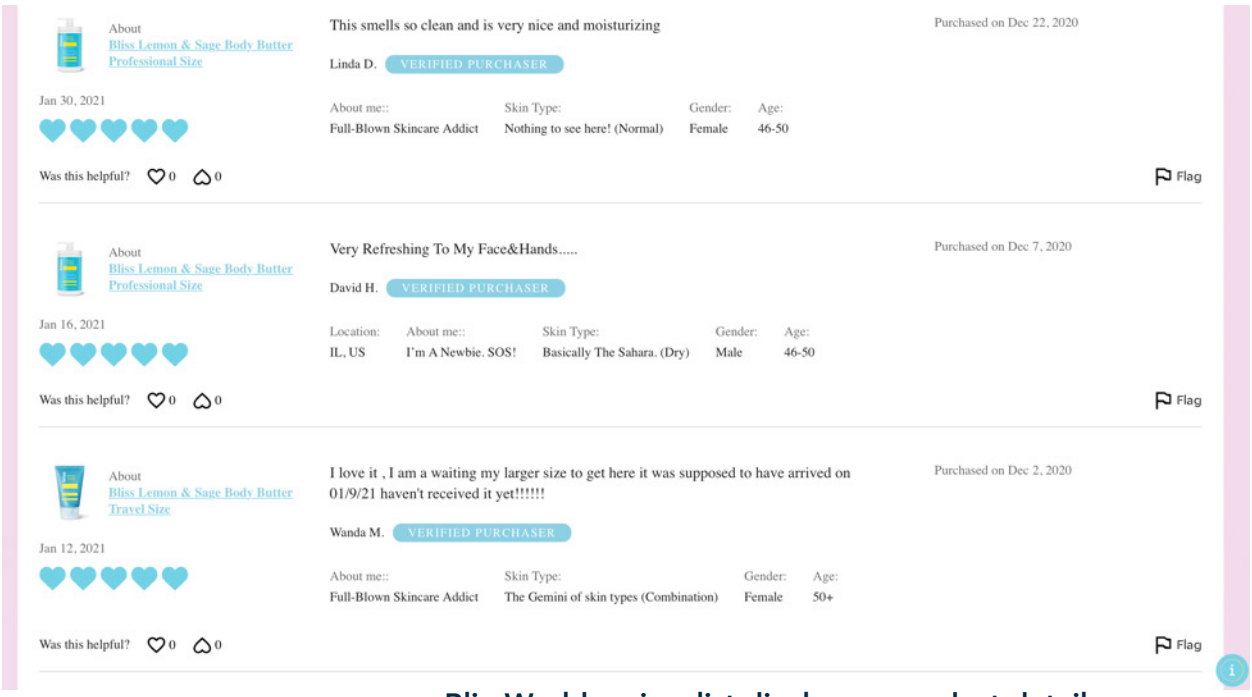
Nike individual review display on the product detail page

Since the main purpose of customer profile attributes is to enable shoppers to find reviews by "people like them", these data are often used in filters. In fact, one way to determine what will be the most useful customer profile attributes to collect is to start with the filters you want to power and work backwards. Rockler Woodworking enables shoppers to filter reviews based on the woodworking experience of the reviewer. In the example below, this enables a shopper who is a beginner woodworker to easily find 8 highly relevant reviews, rather than having to wade through a much longer set of reviews, many written by people not at all like the shopper.



Rockler Woodworking review filters and individual review display

It is also a best practice to show customer profile attributes along with each individual review. That makes is easy for a shopper to see whether the review was written by someone like them or not, when just browsing the review list without a filter. BlissWorld makes these data particularly easy to see by grouping them directly under the main body of the review:



BlissWorld review list display on product detail page

Though often tempting, don't rush to use location as a customer profile attribute. Does the state and city that a reviewer comes from really affect how a shopper feels about the review? If so, go with it; but if the goal is just to humanize the reviewer, you may be better off saving the space for attributes that provide more useful information to the shopper. The same applies to other generic attributes like gender or age. If the gender or age of the reviewer affects the relevance of the review in the eyes of the shopper, then those are good questions to ask. Otherwise, see what alternative questions might be more valuable. One test to see if a customer profile attribute question is worth asking is to think whether it's important enough to use as a filter. Would a shopper care to filter for reviews by people from their area? Would a shopper care to filter for reviews by people of about the same age?

In general, don't skimp on customer profile attributes or assume that rating subdimensions are more important. For many products, the optimal format may be just a core review – stars, title, body – and then a bunch of customer profile attributes to help a shopper determine which reviews to read.

How many subdimensions to collect

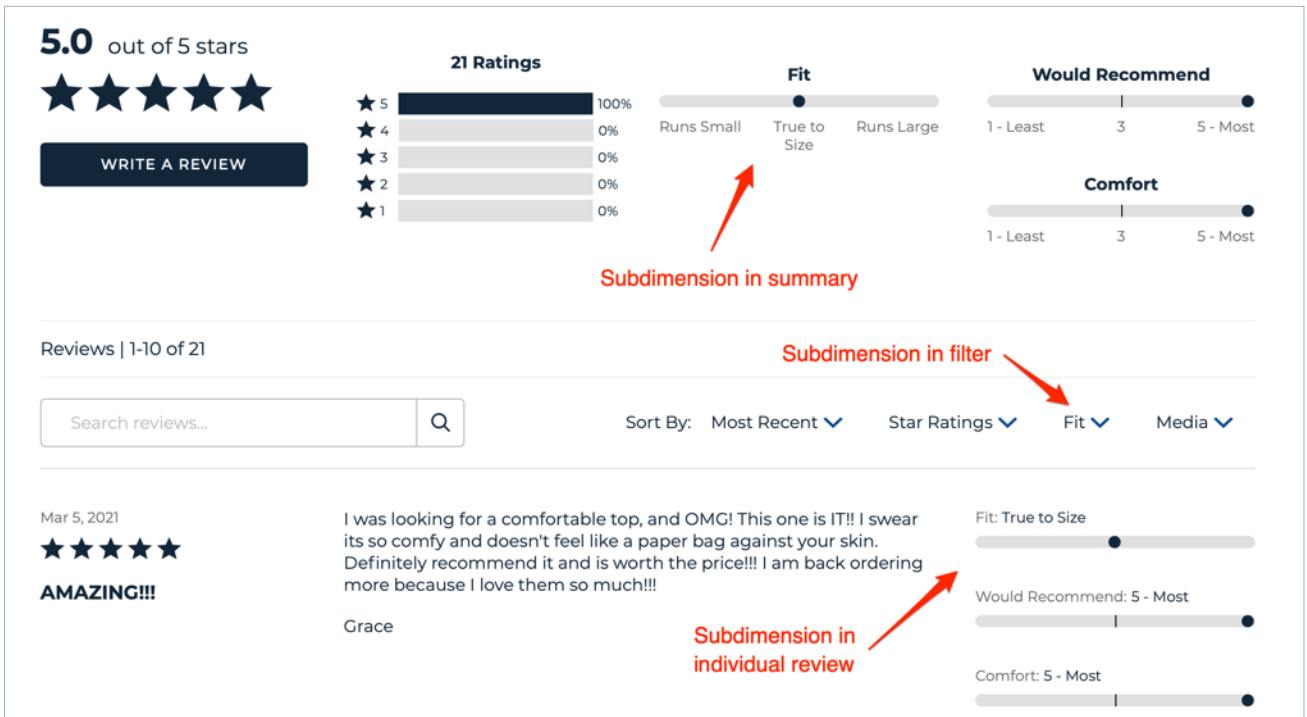
A rule of thumb is that each subdimension you add to the collection form decreases response rates by about 2%. So if you use just a few subdimensions, the total number of reviews you collect won't be affected by much. If you use 10 subdimensions, expect to lose about 20% of your total review volume vs a vanilla stars-title-body review form. This means that if you are choosing your subdimensions wisely, the value you are creating with them is a lot greater than what you are losing by collecting a few less reviews total. But you shouldn't throw in subdimensions that aren't there for a good reason.

However, there is a bit more nuance to the picture. Reviewers get annoyed when subdimension questions seem irrelevant, and they appreciate it when they are well chosen. So the fall-off in review volume is a bit less with "good" subdimension questions and a bit more with "bad" ones. Also, the usual rule about the danger of submission friction applies: highly motivated reviewers make it through even long forms, while lightly motivated reviewers are more likely to abandon. And angry/disappointed reviewers tend to be more highly motivated. Therefore, the friction of long forms creates a small negative bias in the overall sentiment of reviews that are submitted, since the fall off in positive-sentiment reviews tends to be a smidge higher than the fall off in negative-sentiment reviews. That alone should not deter you from using subdimensions; it should just encourage you to think carefully in constructing high quality subdimensions.

HOW TO DISPLAY

Using subdimensions in the summary, with individual reviews, and in filters.

As described above, subdimensions are generally used in three areas of the reviews display: the summary, the list of individual reviews, and the filters.



Scrubs and Beyond reviews summary and list on product detail page

Not all subdimensions should be used in all areas. The usual rules about avoiding clutter and less-can-be-more apply.

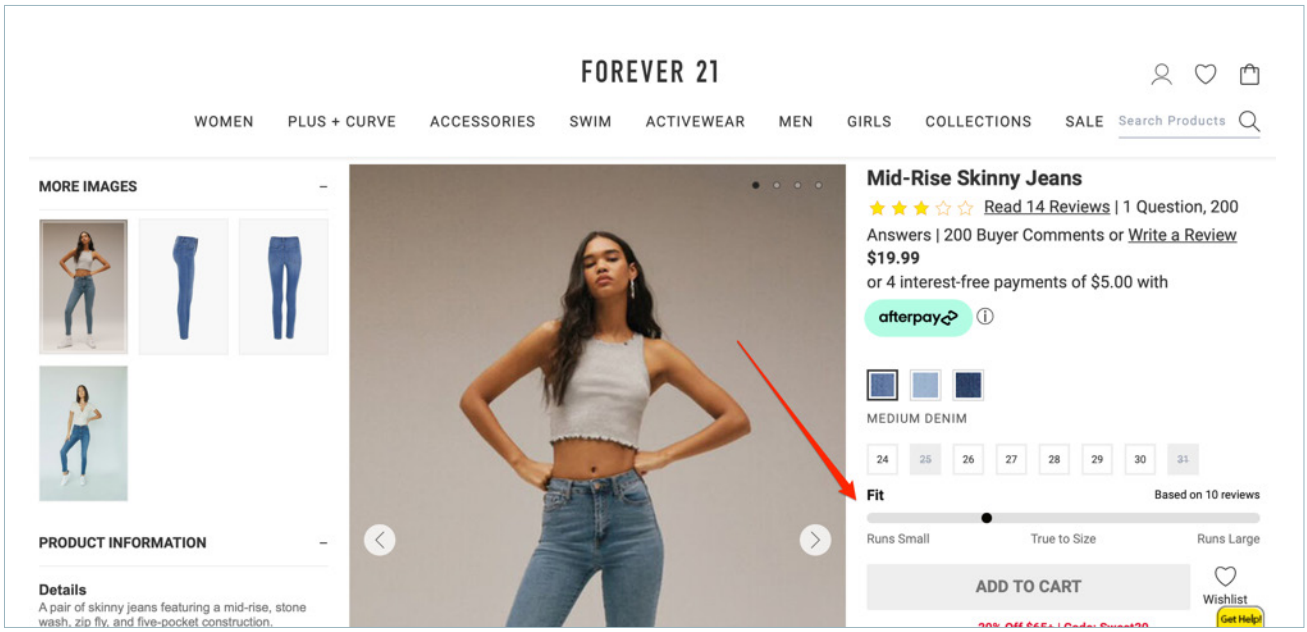
THE GENERAL RULES			
Type	**Summary**	**Individual Reviews**	**Filters**
Describes item performance	Yes! Show averages and distributions.	Maybe. Is the data point valuable in isolation, or only in aggregate?	Maybe. Is it useful to the shopper to see just those reviews where the reviewer said this?
Describes reviewer	Maybe. Is it helpful for a shopper to know stats on the population of reviewers?	Yes! Help shoppers see the ways in which the reviewer is like them or different.	Yes! Make it easy for shoppers to find reviews by "people like them".

Quantitative and evaluative rating subdimensions are useful primarily to provide an aggregate view of item performance across a body of reviews. Therefore, they are typically shown in the review summary area, using an average or a distribution. To determine whether to show rating subdimensions along with the individual reviews, ask whether the data are useful individually or only in aggregate. Is it helpful to the shopper to see that Bob thought the product's battery lasted a bit longer than advertised? Or is it only helpful to know that the average of 324 people's opinions is that the battery lasted a bit longer than average?

If the former, include the data point along with the individual review. If the latter, save the space in the list view of individual reviews and only show the subdimension in the summary. To determine whether a rating subdimension is useful as a filter or not, ask whether a shopper would ever want to see just reviews by people who chose a certain value of that subdimension.

Customer profile attributes and some rating subdimensions (like those about usage or variant purchased) are the opposite. Since they don't say anything about the item itself, they often don't work well in the review summary. However, by enabling shoppers to easily see reviews by "people like them", these data allow shoppers to take sets of reviews that are much too large to be read and filter to easily see just those which are most valuable.

Pro tip: For apparel and footwear, put the unit showing real-world fit next to the size selector. Not all shoppers make it down to the reviews area of your product page, and if your items run small or large, you want to be sure shoppers know that before they purchase the wrong size and you end up with a return. Forever 21 provides a great example of this:



Configuring filters

Once you have determined which subdimensions should be used as filters, you have a few more decisions: What order should the values in each filter appear, should you show counts of the reviews that will result from applying each filter value, and should you show filter values which would result in zero reviews appearing?

Order

There are two choices here: show the values in the order in which they appeared on the collection form or sort them from most-to-least frequently selected. In general, use the input order when there is a logic to the sequence, for example, sizes from largest to smallest, time from longest to shortest, or any sort of value scale (like 5 stars down to 1 star). Choose frequency of selection as the sort order for pick lists where there is no inherent logic to the sequence. That helps shoppers understand how many results to expect for any filter they choose.

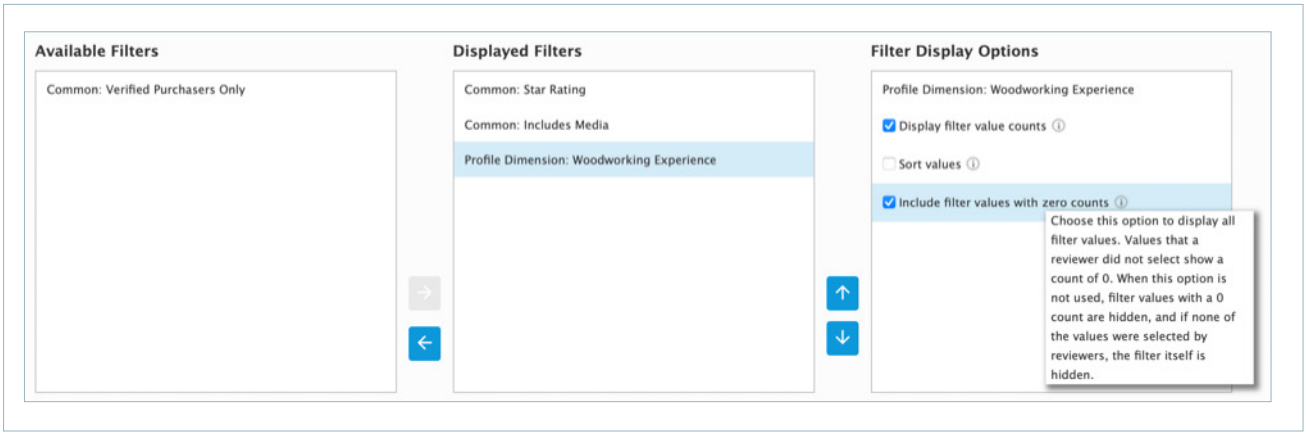
Counts

In general, it is helpful for shoppers to know how many results to expect when applying a filter. There is no point in applying filters that will produce an empty result set. Also, filter counts can be useful to help a shopper see the distribution of selections made by reviewers. For example, if the filter shows the primary use for which the item was bought, a shopper can see if their intended use is common or uncommon. If the filter shows skin type, the shopper can infer what skin type most other buyers of the product have. So most of the time you will want to display the counts. However, if you are concerned that showing counts might bias a shopper away from a purchase, then hide the counts. (e.g. "No one like me is buying this, maybe I shouldn't either…").

Showing zero count values

Should you show filter values which have no associated reviews? There's no use to a shopper in filtering by these values, so you can hide them in order to "clean up" the list. However, you may want to show these values anyway, if you feel that there is significance to the zero count. It probably does not mean much if there are some zero count values in a size-purchased filter – it just happens that no one who bought that size wrote a review. But with an intended-use filter, it may be meaningful to the shopper that no one chose a particular use while many chose others – it implies that use is an outlier.

Here's an example of the configuration used by Rockler Woodworking for the expertise filter shown above:



INTERNAL USE OF SUBDIMENSIONS

Thus far, we have focused only on the use of subdimensions by shoppers. However, you may find subdimensions are highly useful for your internal review analytics, too. A review is, after all, just a survey form; you can change the questions you ask with ease, and, most important, you don't need to display all the subdimensions you collect.

A good practice is to use the same thought process you used in designing subdimensions for shoppers, but instead think about an internal audience. What topics do you want to be sure you get feedback on that might not be consistently addressed in the review text? What profile information on reviewers will help you to better understand their viewpoints?

Maybe there are particular aspects of a new product line that you want to drill down on. How do owners feel about that new feature you added? Is the item holding up as well under real-world use as you promised? Are there particular expectations customers had for the product that weren't met? You may notice a pattern here: you can use unpublished questions to probe for weaknesses that are important for you to know about but that you might not want to publish to your product pages.

You can also use customer profile attributes to gain advanced insights by correlating product performance metrics with profile attributes. We refer to this as "de-averaging". For example, the average overall rating for a pair of jeans might be 4.2 stars. But if you collect body type as a customer profile attribute, you can break that apart. Perhaps reviewers with body type "curvy" gave it an average score of 4.5 while those with body type "athletic" gave it a 3.9. This analysis could be refined further by collecting a rating subdimension on comfort and de-averaging that against body type. There's plenty of low-hanging fruit in basic demographic segmentation. Do your male customers rate it as highly as your female customers? How about your older customers vs your younger ones? But don't stop with just the overall rating. Maybe that 4.0 ease-of-assembly score is really a composite of younger customers who rated that aspect a 4.8 and older customers who rated it a 3.2. Maybe, in looking at reviews for your range of basketball shoes, you'll discover that among 3 shoes with the same overall rating, one shoe has higher ratings from those who play on indoor courts, while another shoe gets top marks from those who play outdoors. Insights like that can be of great value to product development teams and to marketing and merchandising.

Another advanced use of customer profile attributes is to collect data to augment your CRM or loyalty systems. It's a small percent of your total customer base that write reviews, so you should not expect this strategy to transform the overall data quality of your house file. However, the customers that write reviews are typically your most engaged and may be those you want to connect with to become ambassadors, or at least invite to higher levels of your loyalty program. Using subdimensions (perhaps unpublished) on the review form can be an easy way to collect selected bits of demographic or profile data to tune these programs.

SUMMARY

User testing has shown that most shoppers only read a handful of reviews. This means that if you have collected a healthy number of reviews about your products, most are going unread. Subdimensions are a great tool to help you get full value from all that extra content you are collecting by effectively summarizing key sentiment and by helping shoppers find the most relevant reviews to read. Just follow this formula:

1. What topics do you want to be sure reviewers address to help shoppers make confident purchase decisions? Collect rating subdimensions for these. Where possible, design them as comparisons to benchmarks or use quantified metrics. Aggregate the data in the review summary using averages or distribution counts.

2. What information would help shoppers find the particular reviews that will be most relevant to them? Consider the profile of the reviewer and also their intended use of the item or the particular variant bought. Collect subdimensions for these and use them to power filters.

3. How could structured data increase the insights your team gets from your product reviews? Add additional subdimensions if you are not already capturing what you need – you don't have to display these to shoppers. Set up processes to deliver that information into the right parts of the organization.

Way to go, subdimensions!